

25.11.03

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

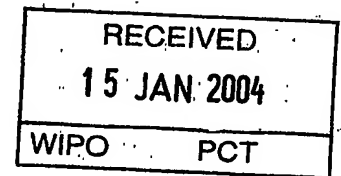
別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日  
Date of Application: 2003年 5月26日

出 願 番 号  
Application Number: 特願2003-148242

[ST. 10/C]: [JP2003-148242]



出 願 人  
Applicant(s):

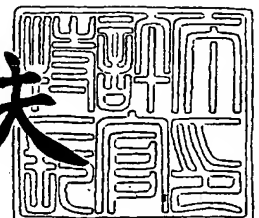
東レ株式会社  
財団法人 東京都医学研究機構  
ヨハネス グーテンベルク ユニベルシティート マインツ

PRIORITY DOCUMENT  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH  
RULE 17.1(a) OR (b)

2003年12月26日

特許庁長官  
Commissioner,  
Japan Patent Office

今 井 康 夫



【書類名】 特許願

【整理番号】 P03-0429

【特記事項】 特許法第30条第1項の規定の適用を受けようとする特許出願

【提出日】 平成15年 5月26日

【あて先】 特許庁長官 殿

【国際特許分類】 C12N 15/00

【発明の名称】 遺伝子型 2 a の C 型肝炎ウイルス (HCV) ゲノム由来の核酸を含む核酸構築物、及び該核酸構築物を導入した細胞

【請求項の数】 21

【発明者】

【住所又は居所】 東京都板橋区成増 3-37-1-302

【氏名】 脇田 隆字

【発明者】

【住所又は居所】 愛知県名古屋市瑞穂区松月町 1-14-206

【氏名】 加藤 孝宣

【発明者】

【住所又は居所】 神奈川県川崎市中原区新城 3-13-5-303

【氏名】 伊達 朋子

【特許出願人】

【識別番号】 000003159

【氏名又は名称】 東レ株式会社

【特許出願人】

【識別番号】 591063394

【氏名又は名称】 財団法人 東京都医学研究機構

## 【特許出願人】

【住所又は居所】 ドイツ連邦共和国 55099 マインツ

【住所又は居所原語表記】 55099 Mainz Germany

【氏名又は名称】 ヨハネス グーテンベルク ユニベルシテート マイ  
ンツ

【氏名又は名称原語表記】 Johannes Gutenberg-Universitaet Mainz

## 【代理人】

【識別番号】 100091096

## 【弁理士】

【氏名又は名称】 平木 祐輔

## 【選任した代理人】

【識別番号】 100118773

## 【弁理士】

【氏名又は名称】 藤田 節

## 【選任した代理人】

【識別番号】 100096183

## 【弁理士】

【氏名又は名称】 石井 貞次

## 【手数料の表示】

【予納台帳番号】 015244

【納付金額】 21,000円

## 【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 遺伝子型2aのC型肝炎ウイルス（HCV）ゲノム由来の核酸を含む核酸構築物、及び該核酸構築物を導入した細胞

【特許請求の範囲】

【請求項1】 遺伝子型2aのC型肝炎ウイルスのゲノムRNA上の、5'非翻訳領域と、NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列と、3'非翻訳領域とを少なくとも含む塩基配列からなる、レプリコンRNA。

【請求項2】 少なくとも1つの選択マーカー遺伝子若しくはリポーター遺伝子及び少なくとも1つのIRES配列をさらに含む、請求項1記載のレプリコンRNA。

【請求項3】 配列番号9又は10で示される塩基配列からなる5'非翻訳領域と、少なくとも1つの選択マーカー遺伝子若しくはリポーター遺伝子と、IRES配列と、遺伝子型2aのC型肝炎ウイルスのゲノムRNA上のNS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列と、配列番号11又は12で示される塩基配列からなる3'非翻訳領域とを含む塩基配列からなる、レプリコンRNA。

【請求項4】 遺伝子型2aのC型肝炎ウイルスのゲノムRNAが、配列番号3又は5で示される塩基配列からなるRNAである、請求項1～3のいずれか1項記載のレプリコンRNA。

【請求項5】 以下の(a)又は(b)のRNAからなるレプリコンRNA。

(a) 配列番号1又は2で示される塩基配列からなるRNA

(b) 配列番号1又は2で示される塩基配列において1～10個の塩基が欠失、置換又は付加された塩基配列からなり、かつ、自律複製能を有するRNA

【請求項6】 請求項1～5のいずれか1項記載のレプリコンRNAを細胞に導入することにより作製された、レプリコン複製細胞。

【請求項7】 細胞が真核細胞である、請求項6記載のレプリコン複製細胞。

【請求項8】 真核細胞がヒト肝由来細胞である、請求項7記載のレプリコ



ン複製細胞。

【請求項 9】 真核細胞がHuh7細胞である、請求項 7 記載のレプリコン複製細胞。

【請求項 10】 C型肝炎ウイルス感染の治療のための治療剤または診断剤の製造または評価のための、請求項 1～5 のいずれか 1 項記載のレプリコン RNA。

【請求項 11】 C型肝炎ウイルス感染の治療のための治療剤または診断剤の製造または評価のための、請求項 6～9 のいずれか 1 項記載のレプリコン複製細胞。

【請求項 12】 C型肝炎ウイルス感染に対するワクチンの製造のための、請求項 1～5 のいずれか 1 項記載のレプリコン RNA。

【請求項 13】 C型肝炎ウイルス感染に対するワクチンの製造のための、請求項 6～9 のいずれか 1 項記載のレプリコン複製細胞。

【請求項 14】 請求項 6～9 のいずれか 1 項記載のレプリコン複製細胞からレプリコン RNA を抽出することを含む、遺伝子型 2a の C型肝炎ウイルスのレプリコン RNA の製造方法。

【請求項 15】 請求項 6～9 のいずれか 1 項記載のレプリコン複製細胞を培養し、得られる培養物からウイルスタンパク質を取得することを含む、遺伝子型 2a の C型肝炎ウイルスのウイルスタンパク質の製造方法。

【請求項 16】 被験物質の存在下で、請求項 6～9 のいずれか 1 項記載のレプリコン複製細胞を培養し、得られる培養物中のレプリコン RNA の複製を検出することを含む、C型肝炎ウイルスの複製を促進又は抑制する物質をスクリーニングする方法。

【請求項 17】 請求項 6～9 のいずれか 1 項記載のレプリコン複製細胞から複製レプリコン RNA を取得し、取得した複製レプリコン RNA を該レプリコン複製細胞とは別の細胞に導入して新たなレプリコン複製細胞を作製する工程を 1 回以上行うことを含む、遺伝子型 2a の C型肝炎ウイルスのレプリコン RNA に複製効率を増大させる突然変異を導入する方法。

【請求項 18】 複製効率の増大が、レプリコン複製細胞に最初に導入され

たレプリコンRNAの複製効率と比較して、少なくとも2倍の増大である、請求項17記載の方法。

【請求項19】 請求項6～9のいずれか1項記載のレプリコン複製細胞から複製レプリコンRNAを取得し、取得した複製レプリコンRNAを該レプリコン複製細胞とは別の細胞に導入して新たなレプリコン複製細胞を作製する工程を1回以上行うこと、及び最終的に得られたレプリコン複製細胞から複製レプリコンRNAを取得することを含む、複製効率が増大した遺伝子型2aのC型肝炎ウイルスのレプリコンRNAを製造する方法。

【請求項20】 請求項19に記載の方法によって製造された複製効率が増大したレプリコンRNAについて、レプリコン複製細胞に最初に導入されたレプリコンRNAとの間の塩基変異又はアミノ酸変異を検出すること、及び複製効率を増大させようとするレプリコンRNAにその検出された塩基変異又はアミノ酸変異を導入することを含む、複製効率が増大した遺伝子型2aのC型肝炎ウイルスのレプリコンRNAを製造する方法。

【請求項21】 配列番号1で示される塩基配列上において、以下の(a)～(h):

- (a) 塩基番号7157の部位におけるAからGへの変異、
  - (b) 塩基番号4955の部位におけるCからUへの変異、
  - (c) 塩基番号4936の部位におけるAからGへの変異、
  - (d) 塩基番号5000の部位におけるAからGへの変異、
  - (e) 塩基番号7288の部位におけるAからGへの変異、
  - (f) 塩基番号5901の部位におけるGからUへの変異、
  - (g) 塩基番号6113の部位におけるAからUへの変異、及び
  - (h) 塩基番号2890の部位におけるAからGへの変異、
- からなる群より選択される少なくとも1つの変異を有する塩基配列からなるレプリコンRNA。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、遺伝子型2aのC型肝炎ウイルスのレプリコンRNA、該レプリコンRNAを導入したレプリコン複製細胞、及び該レプリコンRNAの複製効率を増大させる方法に関する。

#### 【0002】

##### 【従来の技術】

C型肝炎ウイルス (Hepatitis C virus、HCV) は、フラビウイルス科に属する、一本鎖の(+)鎖センスRNAをゲノムとするウイルスであり、C型肝炎の原因となることが知られている。近年の研究により、C型肝炎ウイルスは遺伝子型又は血清型により多数の型に分類されることが分かってきた。現在主流であるHCV遺伝子型の分類法である、SimmondsらによるHCV株の塩基配列を用いた系統解析法ではHCVは遺伝子型1a、遺伝子型1b、遺伝子型2a、遺伝子型2b、遺伝子型3a、遺伝子型3bの6タイプに分類され(非特許文献1)、さらにそれらの各タイプがいくつかのサブタイプに分類されている。現在では、HCVの複数の遺伝子型についてゲノム全長の塩基配列が決定されている(特許文献1及び非特許文献2～5)。

#### 【0003】

HCVは持続的に感染することにより慢性肝炎を引き起こす。現在、世界的規模で認められる慢性肝炎の主たる原因がHCV持続感染である。実際、持続感染者の50%程度が慢性肝炎を発症し、そのうち約20%の患者が10年～20年を経て肝硬変に移行し、さらにその一部は肝癌といった致死的な病態へと進展する。

#### 【0004】

C型肝炎に対する現在の主な治療は、インターフェロン- $\alpha$ 、インターフェロン- $\beta$ 、及びインターフェロン- $\alpha$ とプリンヌクレオシド誘導体であるリバビリンとの併用療法により行われている。しかしながら、これらの治療を行っても、全治療者の約60%に治療効果が認められるだけであり、効果が出た後に治療を中止すると半分以上の患者が再燃する。インターフェロンの治療効果は、HCVの遺伝子型と関連することが知られており、遺伝子型1bに対しては効果が低く、遺伝子型2aに対してはより効果が高いと言われている(非特許文献6)。

## 【0005】

工業国において罹患率が高く、最終的に深刻な結果を招き、かつ現在は原因治療が存在しないC型肝炎に対する効果的な治療薬又は予防薬の開発は重要な目標である。そのため、HCV特異的な化学療法、ワクチン療法の発展が切望されている。抗HCV薬開発のターゲットとしては、HCVの複製抑制やHCVの細胞感染の抑制が考えられる。

## 【0006】

最近まで、HCVを細胞培養系で増やすこと、培養細胞に感染させることは困難であり、また、HCVに感染可能かつ実験可能な動物はチンパンジーしかなかったため、HCVの複製機構や感染機構の研究は困難であった。しかし最近になって、HCV由来の自律複製能を有するRNAとして、HCVサブゲノムRNAレプリコンが作製されたことにより（特許文献2、非特許文献7～10）、培養細胞を用いてHCVの複製機構を解析することが可能となった。これらのHCVサブゲノムRNAレプリコンは、遺伝子型1bのHCVゲノムRNAの5'非翻訳領域中のHCV IRESの下流に存在する構造タンパク質を、ネオマイシン耐性遺伝子及びその下流に連結したEMCV IRESによって置換したものである。このRNAレプリコンは、ヒト肝癌細胞Huh7に導入してネオマイシン存在下で培養することにより、Huh7細胞内で自律複製することが証明された。

## 【0007】

しかしながら、このようなHCVの細胞内RNA複製系は、未だ遺伝子型1bのHCVのゲノムRNAを用いたものしか作製されていない。異なる遺伝子型のHCVではコードされるウイルスタンパク質にも違いがあることが報告されていることから、遺伝子型1bのHCV由来のサブゲノムRNAレプリコンの解析だけでは、HCVの複製機構を十分に解明することは難しいと考えられる。さらに、インターフェロンの治療効果がHCVの遺伝子型によって異なることから、遺伝子型1bのHCVのサブゲノムRNAレプリコンを含むHCV複製系のみを用いて色々なタイプのHCVに効果を及ぼす抗HCV薬を開発することは特に難しいと考えられる。

## 【0008】

## 【特許文献 1】

特開 2 0 0 2 - 1 7 1 9 7 8 号公報

## 【特許文献 2】

特開 2 0 0 1 - 1 7 1 8 7 号公報

## 【非特許文献 1】

Simmonds, P. et al, Hepatology, (1994) 10, p. 1321-1324

## 【非特許文献 2】

Choo et al., Science, (1989) 244, p. 359-362

## 【非特許文献 3】

Kato et al., J. Med. Virol., (2001) 64(3) p. 334-339

## 【非特許文献 4】

Okamoto, H et al, J. Gen. Virol., (1992) 73 p. 673-679

## 【非特許文献 5】

Mori, S. et al, Biochem. Biophys. Res. Commun., (1992) 183, p.  
334-342

## 【非特許文献 6】

Yoshioka et al., Hepatology, (1992) 16(2): p. 293-299

## 【非特許文献 7】

Lohmann et al., Science, (1999) 285, p. 110-113

## 【非特許文献 8】

Blight et al., Science, (2000) 290, p. 1972-1974

## 【非特許文献 9】

Friebe et al., J. Virol., (2001) 75(24): p. 12047-12057

## 【非特許文献 10】

Ikeda et al., J. Virol., (2002) 76(6): p. 2997-3006

## 【0 0 0 9】

## 【発明が解決しようとする課題】

本発明は、未だレプリコン RNA が作製されていない遺伝子型の HCV 由来の  
レプリコン RNA を提供することを目的とする。

## 【0010】

## 【課題を解決するための手段】

本発明者らは、上記課題を解決すべく鋭意研究を行った結果、遺伝子型2aのHCVのレプリコンRNAを作製することに成功した。

## 【0011】

すなわち、本発明は以下のとおりである。

[1] 遺伝子型2aのC型肝炎ウイルスのゲノムRNA上の、5'非翻訳領域と、NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列と、3'非翻訳領域とを少なくとも含む塩基配列からなる、レプリコンRNA。このレプリコンRNAは、少なくとも1つの選択マーカー遺伝子若しくはリポーター遺伝子及び少なくとも1つのIRES配列をさらに含むことが好ましい。

[2] 配列番号9又は10のいずれか1つで示される塩基配列からなる5'非翻訳領域と、少なくとも1つの選択マーカー遺伝子若しくはリポーター遺伝子と、IRES配列と、遺伝子型2aのC型肝炎ウイルスのゲノムRNA上のNS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列と、配列番号11又は12のいずれか1つで示される塩基配列からなる3'非翻訳領域とを含む塩基配列からなる、レプリコンRNA。

[3] 遺伝子型2aのC型肝炎ウイルスのゲノムRNAが、配列番号3又は5で示される塩基配列からなるRNAである、上記[1]又は[2]記載のレプリコンRNA。

[4] 以下の(a)又は(b)のRNAからなるレプリコンRNA。

(a) 配列番号1又は2で示される塩基配列からなるRNA

(b) 配列番号1又は2で示される塩基配列において1～10個の塩基が欠失、置換又は付加された塩基配列からなり、かつ、自律複製能を有するRNA

[5] 上記[1]～[4]記載のいずれかのレプリコンRNAを細胞に導入することにより作製された、レプリコン複製細胞。このレプリコン複製細胞において、レプリコンRNAを導入する細胞は、真核細胞であることが好ましく、ヒト肝由来細胞であることがより好ましく、Huh7細胞であることがさらに好ましい。

- [6] C型肝炎ウイルス感染の治療のための治療剤若しくは診断剤の製造又は評価のための、上記[1]～[4]記載のレプリコンRNA。
- [7] C型肝炎ウイルス感染の治療のための治療剤若しくは診断剤の製造又は評価のための、上記[5]記載のレプリコン複製細胞。
- [8] C型肝炎ウイルス感染に対するワクチンの製造のための、上記[1]～[4]記載のレプリコンRNA。
- [9] C型肝炎ウイルス感染に対するワクチンの製造のための、上記[5]記載のレプリコン複製細胞。
- [10] 上記[5]記載のレプリコン複製細胞からレプリコンRNAを抽出することを含む、遺伝子型2aのC型肝炎ウイルスのレプリコンRNAの製造方法。
- [11] 上記[5]記載のレプリコン複製細胞を培養し、得られる培養物からウイルスタンパク質を取得することを含む、遺伝子型2aのC型肝炎ウイルスのウイルスタンパク質の製造方法。
- [12] 被験物質の存在下で、上記[5]記載のレプリコン複製細胞を培養し、得られる培養物中のレプリコンRNAの複製を検出することを含む、C型肝炎ウイルスの複製を促進又は抑制する物質をスクリーニングする方法。
- [13] 上記[5]記載のレプリコン複製細胞から複製レプリコンRNAを取得し、取得した複製レプリコンRNAを該レプリコン複製細胞とは別の細胞に導入して新たなレプリコン複製細胞を作製する工程を1回以上行うことを含む、遺伝子型2aのC型肝炎ウイルスのレプリコンRNAに複製効率を増大させる突然変異を導入する方法。この方法においては、複製効率の増大が、レプリコン複製細胞に最初に導入されたレプリコンRNAの複製効率と比較して、少なくとも2倍の増大であることがより好ましい。
- [14] 上記[5]記載のレプリコン複製細胞から複製レプリコンRNAを取得し、取得した複製レプリコンRNAを該レプリコン複製細胞とは別の細胞に導入して新たなレプリコン複製細胞を作製する工程を1回以上行うこと、及び最終的に得られたレプリコン複製細胞から複製レプリコンRNAを取得することを含む、複製効率が増大した遺伝子型2aのC型肝炎ウイルスのレプリコンRNAを製造する方法。

[15] 上記[14]記載の方法によって製造された複製効率が増大したレプリコンRNAについて、レプリコン複製細胞に最初に導入されたレプリコンRNAとの間の塩基変異又はアミノ酸変異を検出すること、及び複製効率を増大させようとするレプリコンRNAにその検出された塩基変異又はアミノ酸変異を導入することを含む、複製効率が増大した遺伝子型2aのC型肝炎ウイルスのレプリコンRNAを製造する方法。

[16] 配列番号1で示される塩基配列上において、以下の(a)～(h):

- (a) 塩基番号7157の部位におけるAからGへの変異、
- (b) 塩基番号4955の部位におけるCからUへの変異、
- (c) 塩基番号4936の部位におけるAからGへの変異、
- (d) 塩基番号5000の部位におけるAからGへの変異、
- (e) 塩基番号7288の部位におけるAからGへの変異、
- (f) 塩基番号5901の部位におけるGからUへの変異、
- (g) 塩基番号6113の部位におけるAからUへの変異、及び
- (h) 塩基番号2890の部位におけるAからGへの変異、

からなる群より選択される少なくとも1つの変異を有する塩基配列からなるレプリコンRNA。

【0012】

【発明の実施の形態】

以下、本発明を詳細に説明する。

#### 1. 本発明に係るHCV由来のレプリコンRNA

C型肝炎ウイルス(HCV)のゲノムは、約9600ヌクレオチドからなる(+)鎖の一本鎖RNAである。このゲノムRNAは、5'非翻訳領域(5'NTR又は5'UTRとも表記する)、構造領域と非構造領域とから構成される翻訳領域、及び3'非翻訳領域(3'NTR又は3'UTRとも表記する)からなる。その構造領域にはHCVの構造タンパク質がコードされており、非構造領域には複数の非構造タンパク質がコードされている。

【0013】

このようなHCVの構造タンパク質と非構造タンパク質は、翻訳領域から一統



きのポリプロテインとして翻訳された後、プロテアーゼによって限定分解を受けて構造タンパク質 (Core、E1、及びE2) と非構造タンパク質 (NS2、NS3、NS4A、NS4B、NS5A、及びNS5B) とが各タンパク質として遊離することにより、生成される。これらの構造タンパク質及び非構造タンパク質 (すなわち、H C V のウイルスタンパク質) のうち、Coreはコアタンパク質であり、E1及びE2はエンベロープタンパク質であり、非構造タンパク質 (NS2、NS3、NS4A、NS4B、NS5A、及びNS5B) はウイルス自身の複製に関与するタンパク質である。NS2はメタロプロテアーゼ活性、NS3はセリンプロテアーゼ活性 (N末端側の 3 分の 1) とヘリカーゼ活性 (C末端側の 3 分の 2) を有することが知られている。またNS4AはNS3のプロテアーゼ活性に対するコファクターであり、NS5BはRNA依存RNAポリメラーゼ活性を有することも報告されている。そして、遺伝子型2aのH C V のゲノムも同様の遺伝子構造を有することがすでに報告されている (特許文献 1) 。

#### 【 0 0 1 4 】

本発明者らは、このような遺伝子型2aのH C V ゲノムを用いて、自律的に複製することが可能なRNAを構築した。すなわち本発明のH C V 由来のレプリコンRNAは、遺伝子型2aのH C V ゲノムの全体又は部分RNAを含む自律複製能を有するRNA構築物である。

#### 【 0 0 1 5 】

本明細書では、自律複製能を有しておりH C V ウイルスゲノムを改変して作製されたRNAを、「レプリコンRNA」又は「RNAレプリコン」と呼び、遺伝子型2aのH C V から人為的に作製される自律複製能を有するRNAを、遺伝子型2aのH C V 由来のレプリコンRNAと称する。本明細書においてH C V 由来のレプリコンRNAは、HCV-RNAレプリコンとも称する。

#### 【 0 0 1 6 】

本発明において、「遺伝子型2aのC型肝炎ウイルス」「遺伝子型2aのH C V」とは、Simmondsらによる国際分類に従って遺伝子型2aと同定されるC型肝炎ウイルスを意味する。本発明における「遺伝子型2aのC型肝炎ウイルス」「遺伝子型2aのH C V」には、天然由来のH C V ゲノムRNAを有するウイルスだけでなく、天然由来のH C V ゲノム配列に人為的な改変を加えたゲノムRNAを有するウ

イルスも包含する。遺伝子型2aのHCVの具体例としては、JFH-1株及びJCH-1株（特許文献1）等のウイルスが挙げられる。

【0017】

さらに「遺伝子型2aのC型肝炎ウイルスのゲノムRNA」とは、遺伝子型2aのC型肝炎ウイルスの一本鎖の(+)鎖センスRNAからなるゲノムの全領域にわたる塩基配列を有するRNAを意味する。限定するものではないが、遺伝子型2aのC型肝炎ウイルスのゲノムRNAは、好ましくは配列番号3又は5で示される塩基配列からなるRNAである。

【0018】

本願明細書において、「5'非翻訳領域（5'NTR又は5'UTR）」、「NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列」、「Coreタンパク質をコードする配列（Core領域又はC領域）」、「E1タンパク質をコードする配列（E1領域）」、「E2タンパク質をコードする配列（E2領域）」、「NS2タンパク質をコードする配列（NS2領域）」、「NS3タンパク質をコードする配列（NS3領域）」、「NS4Aタンパク質をコードする配列（NS4A領域）」、「NS4Bタンパク質をコードする配列（NS4B領域）」、「NS5Aタンパク質をコードする配列（NS5A領域）」、「NS5Bタンパク質をコードする配列（NS5B領域）」、及び「3'非翻訳領域（3'NTR又は3'UTR）」、並びにその他の特定の領域若しくは部位は、遺伝子型2aのHCVであるJFH-1株のゲノム全領域をコードする全長cDNA（JFH-1クローン）の塩基配列（配列番号3）を基準として、定めるものとする。配列番号3の塩基配列は、国際DNAデータベース（DDBJ/EMBL/GenBank）からアクセス番号AB047639により取得可能である。具体的には、配列番号3で示される塩基配列に対して特定のHCVのRNA配列をアラインメントしたときに、配列番号3で示される塩基配列上の塩基番号1～340にアラインメントされる配列がそのRNAの「5'非翻訳領域」、同塩基番号3431～9442にアラインメントされる配列が「NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列」、同塩基番号3431～5323にアラインメントされる配列が「NS3タンパク質をコードする配列」、塩基番号5324～5485にアラインメントされる配列が「NS4Aタンパク質をコード

する配列」、同塩基番号5486～6268にアラインメントされる配列が「NS4Bタンパク質をコードする配列」、同塩基番号6269～7666にアラインメントされる配列が「NS5Aタンパク質をコードする配列」、塩基番号7667～9442にアラインメントされる配列が「NS5Bタンパク質をコードする配列」、同塩基番号9443～9678にアラインメントされる配列が「3'非翻訳領域」である。また、この場合「アラインメント」される配列にはギャップ、付加、欠失、置換等が存在していてもよい。さらに上記の「特定のHCV」は、限定するものではないが、JFH-1株若しくはJCH-1株又はそれらの誘導体であるウイルス株を包含する。

#### 【0019】

本発明に係るHCV RNA-レプリコンの一つの実施形態は、遺伝子型2aのC型肝炎ウイルスのゲノムRNA上の、5'非翻訳領域と、NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列と、3'非翻訳領域とを少なくとも含む塩基配列からなる、レプリコンRNAである。このレプリコンRNAは、少なくとも1つの選択マーカー遺伝子若しくはリポーター遺伝子、及び少なくとも1つのIRES配列をさらに含んでもよい。さらにこのレプリコンRNAは、遺伝子型2aのC型肝炎ウイルスのゲノムRNA上の、NS3、NS4A、NS4B、NS5A及びNS5Bタンパク質以外のウイルスタンパク質をコードする配列を、含んでもよい。

#### 【0020】

本発明に係るHCV RNA-レプリコンの別の好適な実施形態は、配列番号9又は10で示される塩基配列からなる5'非翻訳領域と、少なくとも1つの選択マーカー遺伝子若しくはリポーター遺伝子と、IRES配列と、遺伝子型2aのC型肝炎ウイルスのゲノムRNA上のNS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列と、配列番号11又は12で示される塩基配列からなる3'非翻訳領域とを含む塩基配列からなる、レプリコンRNAである。ここで配列番号9及び10で示される塩基配列は、それぞれ、本発明に係るレプリコンRNAであるrSGREP-JFH1（配列番号1）及びrSGREP-JCH1（配列番号2）の5'非翻訳領域の配列である。また配列番号11及び12で示される塩基配列は、それぞれ、本発明に係るレプリコンRNAであるrSGREP-JFH1（

配列番号 1) 及び rSGREP-JCH1 (配列番号 2) の 3' 非翻訳領域の配列である。

#### 【0021】

本発明に係る HCV RNA-レプリコンのさらに好ましい 1 つの実施形態は、配列番号 1 又は 2 で示される塩基配列からなる RNA からなるレプリコン RNA である。さらに、この配列番号 1 又は 2 で示される塩基配列において、1～50 個、1～30 個、1～10 個、1～6 個、1～数個 (2～5 個) の塩基が欠失、置換又は付加された塩基配列からなるレプリコン RNA であって、かつ、自律複製能を有する RNA も、好適な実施形態として本発明の範囲に含まれる。本発明において「自律複製能を有する」とは、レプリコン RNA を細胞中に導入したときに、そのレプリコン RNA が細胞内でそのレプリコン RNA 自身の全長を複製することができることを意味する。限定するものではないが、この自律複製能は、例えば、レプリコン RNA を Huh7 細胞中にトランスフェクションし、Huh7 細胞を培養し、得られる培養物中の細胞から抽出した RNA について、導入したレプリコン RNA を特異的に検出可能なプローブを用いたノーザンブロットハイブリダイゼーションを行ってレプリコン RNA の存在を検出することにより、確認することができる。自律複製能を確認するための具体的な操作は、本明細書の実施例に記載されたコロニー形成能の測定、HCV タンパク質の発現確認、レプリコン RNA の検出等の記載に従って行うことができる。

#### 【0022】

本発明において「選択マーカー遺伝子」とは、その遺伝子が発現された細胞だけが選択されるような選択性を細胞に付与することができる遺伝子を意味する。選択マーカー遺伝子の一般的な例としては抗生物質耐性遺伝子が挙げられる。本発明において好適な選択マーカー遺伝子の例としては、ネオマイシン耐性遺伝子、チミジンキナーゼ遺伝子、カナマイシン耐性遺伝子、ピリチアミン耐性遺伝子、アデニリルトランスフェラーゼ遺伝子、ゼオシン耐性遺伝子、ピューロマイシン耐性遺伝子等が挙げられるが、ネオマイシン耐性遺伝子、チミジンキナーゼ遺伝子が好ましく、ネオマイシン耐性遺伝子がさらに好ましい。但し本発明における選択マーカー遺伝子はこれらに限定されるものではない。

#### 【0023】

また本発明において「リポーター遺伝子」とは、その遺伝子発現の指標となる遺伝子産物をコードするマーカー遺伝子を意味する。リポーター遺伝子の一般的な例としては、発光反応や呈色反応を触媒する酵素の構造遺伝子が挙げられる。本発明において好適なリポーター遺伝子の例としては、トランスポゾン Tn9 由来のクロラムフェニコールアセチルトランスフェラーゼ遺伝子、大腸菌由来の  $\beta$  グルクロニダーゼ若しくは  $\beta$  ガラクトシダーゼ遺伝子、ルシフェラーゼ遺伝子、緑色蛍光タンパク質遺伝子、クラゲ由来のイクリオン遺伝子、分泌型胎盤アルカリフォスファターゼ (SEAP) 遺伝子等が挙げられる。但し本発明におけるリポーター遺伝子はこれらに限定されるものではない。

#### 【0024】

上記の選択マーカー遺伝子及びリポーター遺伝子は、レプリコン RNA 中にどちらか一方のみが含まれていてもよいし、両方が含まれていてもよい。

#### 【0025】

本発明における「IRES配列」とは、RNAの内部にリボソームを結合させて翻訳を開始させることが可能な内部リボソーム結合部位を意味する。本発明における IRES配列の好適な例としては、以下に限定するものではないが EMCV IRES (脳心筋炎ウイルスの内部リボソーム結合部位)、FMDV IRES、HCV IRES等が挙げられるが、EMCV IRES及びHCV IRESがより好ましく、EMCV IRESが最も好ましい。

#### 【0026】

さらに本発明に係るレプリコン RNA は、他の HCV 株又は他の遺伝子型の HCV のゲノム RNA 上の配列を含んでもよい。例えば、遺伝子型 1b の HCV ゲノムの断片を含んでもよい。他の HCV 株としては、例えば HCV-1、HCV-H、HC-J1、HCT-18、H77、DK-7、US11、S14、HCT23、HCV-Th、DR1、DR4、HCT27、S18、SW1、DK9、H90、TD-6E1、S9、HCV-BK、T10、DK1、HC-J4、HCV-J、HK3、HK8、HK5、HCV-G3、IND5、IND8、P10、D1、D3、SW2、T3、S45、SA10、US6、HCV-JK1、HCV-JK4、HCV-JK3、HCV-JK2、HCV-JT、HC-J2、HCV-T、HK4、HC-G9、Z1、Bi, S.I.、Cho, J.M.、HCV-J6、T4、T9、US10、HC-J5、T2、HC-J7、DK11、SW3、DK8、T8、HC-J8、S83、HK2、HC-J6、HC-J8、BEBE1、HCV-J6、HCV-J8、HD10-2、BR36-9、S52、S54、S2、BR33-1、HK10、DK12、HCV-TR、BA-1、BA-2、DK13、Z1、Z4、Z6、Z7、HK2

、SA1、SA4、SA5、SA7、SA13、SA6、NZL1、SA30、EG-13、HCV-K3a/650、ED43、E UH1480、EUHK2、Th580、VN235、VN405、VN004、JK049、JK046、JFH-1、JCH-1、J CH-2、JCH-3、JCH-4、JCH-5、JCH-6、J6CF、H77等が挙げられるが、これらに限 定するものではない。

#### 【0027】

本発明に係るレプリコンRNAは、好ましくは、最も5'側に遺伝子型2aのHCVのゲノムRNA上の5'非翻訳領域を、最も3'側に遺伝子型2aのHCVのゲノムRNA上の3'非翻訳領域を有する。選択マーカー遺伝子若しくはリポーター遺伝子は、IRES配列の上流に連結されてもよいし、「NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列」の上流又は下流に連結されてもよいし、「NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列」の中に挿入されてもよい。

#### 【0028】

本発明に係るレプリコンRNAは、より好ましくは、最も5'側に遺伝子型2aのHCVのゲノムRNA上の5'非翻訳領域を有し、それよりも下流に選択マーカー遺伝子若しくはリポーター遺伝子と、IRES配列と、「NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列」とをこの順番で有し、さらに最も3'側に遺伝子型2aのHCVのゲノムRNA上の3'非翻訳領域を有する。

#### 【0029】

本発明に係るレプリコンRNAには、上記したような配列の他に、レプリコンRNAを導入する細胞内で発現させたい任意の外来遺伝子を含むRNAを含んでもよい。外来遺伝子は、5'非翻訳領域の下流に連結してもよいし、選択マーカー遺伝子若しくはリポーター遺伝子の上流又は下流に連結させてもよいし、「NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列」の上流又は下流に連結してもよいし、「NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列」の中に挿入してもよい。外来遺伝子を含むレプリコンRNAは、導

入された細胞内で翻訳される際に、該外来遺伝子にコードされたタンパク質を発現することができる。従って外来遺伝子を含むレプリコンRNAは、遺伝子治療などの、特定の遺伝子産物を細胞内で生成させることを目的とする場合にも、好適に使用することができる。

#### 【0030】

また本発明に係るレプリコンRNAには、さらにリボザイムを含んでいてもよい。リボザイムは、5'側のレプリコンRNA中の選択マーカー遺伝子、リポーター遺伝子又は外来遺伝子と、それより3'側のIRES配列及び「NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列」とを連結するように挿入し、リボザイムの自己切断活性により両者が切断されて分離するようにすることができる。

#### 【0031】

本発明に係るレプリコンRNAにおいては、上述したような選択マーカー遺伝子、リポーター遺伝子、遺伝子型2aのC型肝炎ウイルスのゲノムRNA上のウイルスタンパク質をコードする配列、遺伝子型2a以外のHCVのウイルスタンパク質コード配列、及び外来遺伝子等が、レプリコンRNAから正しい読み枠で翻訳されるように連結される。それらの配列のうちでタンパク質をコードする配列は、遺伝子型2aのC型肝炎ウイルスの「NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列」から翻訳されるポリタンパク質とともに融合タンパク質として発現させた後で、プロテアーゼによって各タンパク質へと分離するように、プロテアーゼ切断部位等を介して互いに連結させてもよい。

#### 【0032】

### 2. 本発明に係るレプリコンRNAの作製

本発明に係るHCV RNA-レプリコンは、当業者に公知である任意の遺伝子工学的手法を用いて作製することができる。限定するものではないが、HCV RNA-レプリコンは、例えば以下のような方法で作製することができる。

#### 【0033】

まず、遺伝子型2aのC型肝炎ウイルスのゲノムRNAの全領域に対応するDN

Aを、常法によりRNAプロモーターの下流に連結し、DNAクローンを作製する。ここで、「RNAに対応するDNA」とは、当該RNAの塩基配列のU（ウラシル）をT（チミン）に置き換えた塩基配列を有するDNAを意味する。前記RNAプロモーターは、プラスミドクローン中に含まれるものであることが好ましい。RNAプロモーターとしては、限定するものではないが、T7 RNAプロモーターが特に好ましい。

#### 【0034】

次に、作製したDNAクローンについて、例えば、5'非翻訳領域の下流に位置する構造領域（Core配列、E1配列、E2配列）及びNS2タンパク質をコードする配列を、選択マーカー遺伝子若しくはレポーター遺伝子とその下流に連結したIRES配列とを含むDNA断片によって置換する。この置換においては、構造領域以外の部分、例えば5'非翻訳領域の3'末端側の断片及びNS3タンパク質をコードする配列の一部が、別の遺伝子型のHCVに由来する配列に置換されてもよい。

#### 【0035】

次いで、その置換したDNAクローンを鋳型として、RNAポリメラーゼによりRNAを合成する。RNA合成は、5'非翻訳領域及びIRES配列から、常法により開始させることができる。鋳型DNAがプラスミドクローンの場合には、そのプラスミドクローンから、RNAプロモーターの下流に連結された上記DNA領域を制限酵素によって切り出して、そのDNA断片を鋳型として用いてRNAを合成することもできる。なお、好ましくは合成されるRNAの3'末端がウイルスゲノムRNAの3'非翻訳領域と一致し、他の配列が付加されたり削除されたりしないことが好ましい。このようにして合成されたRNAが、本発明に係るレプリコンRNAである。

#### 【0036】

### 3. 遺伝子型2aのHCVのレプリコンRNAを導入したレプリコン複製細胞の作製

上記のようにして作製されるレプリコンRNAを、レプリコンRNAを複製させるべき細胞に導入することにより、レプリコンRNAが持続的に複製されている細胞を得ることができる。本明細書では、レプリコンRNAが持続的に増幅さ



れている細胞を「レプリコン複製細胞」と称する。

#### 【0037】

レプリコンRNAを導入する細胞としては、継代培養することが可能な細胞であれば任意の細胞を用いることができるが、真核細胞であることが好ましく、ヒト肝由来細胞であることがより好ましく、Huh7細胞細胞であることがさらに好ましい。これらの細胞は、市販のものを利用してよいし、細胞寄託機関から入手して使用してもよいし、任意の細胞（例えば癌細胞又は幹細胞）から株化した細胞を使用してもよい。

#### 【0038】

レプリコンRNAの細胞内への導入は、当業者には公知の任意の技術を使用し行うことができる。そのような導入法としては、例えば、エレクトロポレーション、パーティクルガン法、リポフェクション法、リン酸カルシウム法、マイクロインジェクション法、DEAEセファロース法等が挙げられるが、エレクトロポレーションによる方法が特に好ましい。

#### 【0039】

レプリコンRNAは、目的のレプリコンRNAを単独で導入してもよいし、他の核酸と混合させたものを導入してもよい。導入するRNA量を一定にしてレプリコンRNAの量を変化させたい場合には、目的のレプリコンRNAを、導入する細胞から抽出したトータル細胞性RNAと混合して、細胞内導入に用いればよい。細胞内導入に用いるレプリコンRNAの量は、使用する導入法に応じて決めればよいが、好ましくは1ピコグラム～100マイクログラム、より好ましくは10ピコグラム～10マイクログラムの量を使用する。

#### 【0040】

細胞内導入のために選択マーカー遺伝子又はリポーター遺伝子を含むレプリコンRNAを用いる場合には、そのレプリコンRNAが導入され持続的に複製している細胞を、選択マーカー遺伝子又はリポーター遺伝子の発現を利用して、選択することができる。具体的には、例えば、そのようなレプリコンRNAの細胞内導入処理を施した細胞を、選択マーカー遺伝子又はリポーター遺伝子の発現により選択可能となる培地において培養すればよい。一例として、レプリコンR

NAにネオマイシン耐性遺伝子が選択マーカー遺伝子として含まれる場合には、そのレプリコンRNAを用いて細胞内導入処理した細胞を培養ディッシュに播種し、16～24時間培養した後に、培養ディッシュにG418（ネオマイシン）を0.05ミリグラム/ミリリットル～3.0ミリグラム/ミリリットルの濃度で添加し、その後、週に2回培養液を交換しながら培養を継続し、播種時から好ましくは10日間～40日間、より好ましくは14日間～28日間培養した後にクリスタルバイオレットで生存細胞を染色することにより、レプリコンRNAが導入され持続的に複製されている細胞を、形成されたコロニーとして選択することができる。

#### 【0041】

形成されたコロニーからは、常法により生存細胞をクローン化し、培養を継続することにより、細胞をクローン化することができる。このようにして得られる目的のレプリコンRNAが持続的に複製されている細胞クローンを、本明細書では「レプリコン複製細胞クローン」と称する。

#### 【0042】

樹立した細胞クローンについては、導入されたレプリコンRNAから該細胞クローン中で複製されているレプリコンRNAの検出、導入されたレプリコンRNA中の選択マーカー遺伝子又はリポーター遺伝子の宿主ゲノムDNAへの組み込みの有無の確認、及びHCVタンパク質の発現の確認を行って、実際に目的のレプリコンRNAが持続的に複製されていることを確認することが好ましい。

#### 【0043】

導入されたレプリコンRNAから該細胞クローン中で複製されたレプリコンRNA（本明細書中では、以下便宜的に、「複製レプリコンRNA」と称する）の検出は、当業者には公知の任意のRNA検出法に従って行えばよいが、例えば、細胞クローンから抽出したトータルRNAについて、導入されたレプリコンRNAに対して特異的なDNA断片をプローブとして用いるノーザンハイブリダイゼーション法を実施することにより検出することができる。

#### 【0044】

また導入されたレプリコンRNA中の選択マーカー遺伝子又はリポーター遺伝子の宿主ゲノムDNAへの組み込みの有無の確認は、限定するものではないが、

例えば、細胞クローンから抽出した宿主ゲノムDNAについて該選択マーカー遺伝子又はリポーター遺伝子の少なくとも一部を増幅するPCRを行い、その増幅産物の有無を確認することによって行うことができる。増幅産物が確認された細胞クローンでは、宿主ゲノム中に選択マーカー遺伝子又はリポーター遺伝子が組み込まれていると判断されることから、レプリコンRNA自体は該細胞内で持続的に複製されていない可能性がある。この場合、レプリコンRNAが持続的に複製されているか否かを、次に示すHCVタンパク質の発現の確認実験によって、確認することができる。

#### 【0045】

HCVタンパク質の発現の確認は、例えば、導入されたレプリコンRNAから発現されるべきHCVタンパク質に対する抗体を、細胞クローンから抽出したタンパク質と反応させることによって行うことができる。この方法は、当業者には公知の任意のタンパク質検出法によって行うことができるが、具体的には例えば、細胞クローンから抽出したタンパク質試料をニトロセルロース膜にブロッティングし、それに対して抗HCVタンパク質抗体（例えば、抗NS3特異的抗体、又はC型肝炎患者から採取した抗血清）を反応させ、さらにその抗HCVタンパク質抗体を検出することによって行うことができる。細胞クローンから抽出したタンパク質中からHCVタンパク質が検出されれば、その細胞クローンは、HCV由来のレプリコンRNAが持続的に複製してHCVタンパク質を発現しているものと判断することができる。

#### 【0046】

以上のようにして、目的のレプリコンRNAを持続的に複製していることが確認された細胞クローン（レプリコン複製細胞クローン）を得ることができる。また本発明においては、このレプリコン複製細胞から、例えばRNAを抽出しその中からレプリコンRNAを電気泳動法により分離する等の当業者には公知の任意の方法により、レプリコンRNAを取得することができる。本発明はそのようなレプリコンRNAの製造方法にも関する。さらに本発明に係るレプリコン複製細胞は、HCVタンパク質を製造するために好適に使用することができる。レプリコン複製細胞からのHCVタンパク質の取得は、当業者であれば常法に従って行

うことができる。具体的には例えば、レプリコン複製細胞を培養し、得られる培養物（培養細胞及び培養培地を含む）から常法によりタンパク質を採取し、さらにそのタンパク質から、抗HCVタンパク質抗体を用いた検出等によりウイルスタンパク質を選択的に得ることにより、遺伝子型2aのC型肝炎ウイルスのウイルスタンパク質を製造することができる。

#### 【0047】

また本発明に係るレプリコン複製細胞が、外来遺伝子を含有するレプリコンRNAを持続的に複製している場合には、そのレプリコン複製細胞を用いて外来遺伝子にコードされるタンパク質を発現させて取得することができる。具体的には例えば、レプリコン複製細胞を培養し、得られる培養物（培養細胞及び培養培地を含む）から常法によりタンパク質を採取し、さらにそのタンパク質から、目的のタンパク質に対する抗体を用いた検出等によりタンパク質を選択的に得ることにより、外来遺伝子にコードされたタンパク質を取得することができる。

#### 【0048】

#### 4. 遺伝子型2aのHCVのレプリコンRNAへの複製効率を増大させる突然変異の導入

本発明に係るレプリコン複製細胞において複製され生成されたレプリコンRNA（複製レプリコンRNA）には、複製効率を向上させる突然変異が頻繁に生ずる。このような突然変異は適合変異であると思われる。

#### 【0049】

本発明では、このことを利用して、本発明に係るレプリコンRNAに複製効率を向上させる突然変異の導入を促進することができる。

#### 【0050】

具体的には、第1のレプリコン複製細胞（好ましくは本発明に係るレプリコンRNAを導入したレプリコン複製細胞）から、第1の複製レプリコンRNAを抽出等により取得し、第1の複製レプリコンRNAをさらに別の細胞に再導入して第2のレプリコン複製細胞を作製するという工程を、1回以上、好ましくは1～10回、より好ましくは1～5回、さらに好ましくは1～2回反復的に行うことにより、レプリコン複製細胞中で、レプリコンRNAに複製効率を増大させる突

然変異を高頻度に導入することができる。

#### 【0051】

複製レプリコンRNAを再導入する細胞としては、任意の細胞を用いることができるが、最初にレプリコンRNAを導入した細胞と同じ生物種由来の細胞であることが好ましく、最初にレプリコンRNAを導入した細胞と同じ生物種由来の同じ組織由来の細胞であることが好ましく、最初にレプリコンRNAを導入した細胞と同じ細胞株の細胞であることがさらに好ましい。

#### 【0052】

従って本発明では、上記の方法を用いて、複製効率を増大させる突然変異を導入したレプリコンRNAを製造することができる。すなわち、まず第1のレプリコン複製細胞（好ましくは本発明に係るレプリコンRNAを導入したレプリコン複製細胞）から、第1の複製レプリコンRNAを抽出等により取得し、さらにこの第1の複製レプリコンRNAをさらに別の細胞に再導入して第2のレプリコン複製細胞を作製する工程を、1回以上、好ましくは1～10回、より好ましくは1～5回、さらに好ましくは1～2回反復的に行った後、この反復工程の最後に得られる最終的なレプリコン複製細胞から、複製レプリコンRNAを抽出等によって取得することにより、複製効率が増大したレプリコンRNAを製造することができる。

#### 【0053】

本発明では、以上のような方法により、レプリコンRNAの複製効率を少なくとも2倍、好ましくは10～100倍、より好ましくは100～10000倍に増大させることができる。

#### 【0054】

このような方法により製造した複製効率が増大したレプリコンRNAについては、逆転写PCRによってcDNAを得てそれを塩基配列決定に供するなどの公知の方法により、塩基配列を決定することが好ましい。さらに、決定された塩基配列又はそれにコードされるアミノ酸配列を、最初に細胞に導入されたレプリコンRNAの塩基配列と比較することにより、適合変異を同定することができる。複製効率を増大させる適合変異としては、特に、レプリコンRNAにコードされた

ウイルスタンパク質のアミノ酸を変異させる非同義置換が好ましい。

#### 【0055】

また本発明は、そのようにして同定した適合変異を、複製効率を増大させようとするレプリコンRNAに常法により導入することによって、複製効率が増大した遺伝子型2aのC型肝炎ウイルスのレプリコンRNAを製造することができる方法も提供する。

#### 【0056】

以上のようにして製造された複製効率が増大したレプリコンRNAは、その方法に使用した細胞中においてレプリコンRNAを大量に製造するために使用することができる。

#### 【0057】

本発明に係るレプリコンRNAの複製効率は、当業者に公知の方法により決定することができるが、例えば次のような方法に従って決定すればよい。たとえばHuh7細胞に0.0001、0.0003、0.001、0.003、0.01、0.03、0.1、0.3、1.0マイクログラムの量のレプリコンRNAをトランスフェクションして、前述の実験手法と同様の方法でG418による選択培養を21日間行った後にコロニー形成数（コロニー数）を測定する。導入したRNA量とコロニー形成数とを比較して容量依存的にコロニー形成が増加するレプリコンRNA導入量の範囲を決定し、その範囲内のコロニー形成数を、導入したRNA量で除算して得られる値を、1マイクログラムあたりのコロニー形成率とする。この計算式は、以下のとおりである。

コロニー形成率 [(Colony forming Unit; CFU)/マイクログラム] = コロニー形成数 [個] / 導入したRNA量 [マイクログラム]

#### 【0058】

こうして算出されたコロニー形成率を、導入したレプリコンRNAの複製効率を示す値とする。すなわち、コロニー形成率が高いほど、そのレプリコンRNAの複製効率は高い。

#### 【0059】

またレプリコンRNAの複製効率は、形成されたコロニー1個あたりの導入したレプリコンRNAのコピー数で示されるコロニー形成能で表すこともできる。

すなわち、以下のような計算式に従って算出することができる：コロニー形成能 = 導入したレプリコンRNAのコピー数 [コピー] / コロニー形成数 [個]。

#### 【0060】

##### 5. 本発明の他の実施形態

本発明に係るレプリコンRNA複製細胞は、例えばC型肝炎ウイルスの複製を促進又は抑制する物質をスクリーニングするための試験系として使用することもできる。具体的には例えば、被験物質の存在下で、レプリコン複製細胞を培養し、得られる培養物中のレプリコンRNAの複製を検出し、その被験物質がレプリコンRNAの複製を促進又は抑制するかどうかを判定することにより、C型肝炎ウイルスの複製を促進又は抑制する物質をスクリーニングすることができる。この場合、得られる培養物中のレプリコンRNAの複製の検出は、レプリコンRNA複製細胞から抽出したRNA中のレプリコンRNAの量又は有無を検出することによるものであってもよいし、培養物中または該培養物に含まれるレプリコンRNA複製細胞中のタンパク質に含まれるHCVタンパク質の量又は有無を検出するものであってもよい。

#### 【0061】

このような本発明に係るレプリコンRNA複製細胞を用いる試験細胞系は、C型肝炎ウイルス感染の治療のための治療剤若しくは診断剤の製造又は評価を目的とすることが考えられる。そのような目的としては、具体的には、以下のような例が挙げられる。

#### 【0062】

(1) 遺伝子型2aのHCVの増殖を抑制する物質の探索。このような物質としては、例えば、直接的若しくは間接的に遺伝子型2aのHCVの増殖に影響を及ぼす有機化合物、あるいは遺伝子型2aのHCVゲノム若しくはその相補鎖の標的配列にハイブリダイズすることによりHCVの増殖若しくはHCVタンパク質の翻訳に直接的又は間接的に影響を及ぼすアンチセンスオリゴヌクレオチド等が挙げられる。

(2) 細胞培養中で抗ウイルス作用を有する各種物質の評価。各種物質としては、

合理的ドラッグデザイン又はハイスループットスクリーニングを用いて得られた物質（例えば単離精製された酵素）等が挙げられる。

(3) 遺伝子型2aのHCVに感染した患者の治療のための、新規攻撃標的の同定。  
例えばHCVウイルス増殖のために重要な役割を果たす宿主細胞性タンパク質を同定するために、本発明に係るレプリコン複製細胞を使用することができる。

(4) HCVウイルスの薬剤等に対する耐性獲得能の評価及び該耐性に関わる変異の同定。

### 【0063】

本発明に係るレプリコンRNA又はレプリコンRNA複製細胞はまた、さらに以下のような目的で使用する事が考えられる。

(5) C型肝炎ウイルス感染の診断薬又は治療薬の開発、製造及び評価のために使用可能な抗原としてのウイルスタンパク質の製造。

(6) C型肝炎ウイルス感染の診断薬又は治療薬の開発、製造及び評価のために使用できるHCVウイルス又はウイルス様粒子を製造するための、ウイルスゲノム複製系。

(7) 遺伝子型2aのHCVに対するワクチンとして使用可能なワクチン抗原の製造。

(8) 遺伝子治療用の外来遺伝子を組み込んで使用する、肝細胞指向性遺伝子ベクターの製造。

### 【0064】

#### 【実施例】

本発明を、以下の実施例及び図面に基づいてさらに具体的に説明する。但し、本発明の技術的範囲はこれら実施例に限定されるものではない。

### 【0065】

#### 【実施例1】 レプリコンRNAの作製

##### (A) 発現ベクターの構築

劇症肝炎の患者から分離したC型肝炎ウイルスJFH-1株（遺伝子型2a）のウイルスゲノム全領域に対応するDNAを、該ウイルス株のゲノム全長cDNAを含むJFH-1クローンから取得して、pUC19プラスミドに挿入したT7 RNAプロモーター



配列の下流に挿入した。このようにして構築されたプラスミドDNAを、以下、pJFH1と称する。同様に、慢性肝炎の患者から分離したC型肝炎ウイルスJCH-1株（遺伝子型2a）のウイルスゲノム全領域に対応するDNAを、該ウイルス株のゲノム全長cDNAを含むJCH-1クローンから取得して、pUC19プラスミドに挿入したT7 RNAプロモーター配列の下流に挿入した。このようにして構築されたプラスミドDNAを、以下、pJCH1と称する。なお、上記JFH-1クローン及びJCH-1クロンの作製については、特許文献1及び非特許文献3に記載されている。またJFH-1クロンの全長cDNAの塩基配列は、国際DNAデータバンク（DDBJ/EMBL/GenBank）のアクセッション番号：AB047639に、JCH-1クロンの全長cDNAの塩基配列はアクセッション番号：AB047640に登録されている。

#### 【0066】

このようにして構築したプラスミドDNA pJFH1及びpJCH1の構造を、図1の上段に示す。「T7」はT7 RNAプロモーター、「G」は、挿入したJFH-1又はJCH-1由来DNAの5'端の上流かつT7 RNAプロモーター配列の3'端の下流に挿入したdGTP、を示す。「5'NTR」から「3'NTR」までがC型肝炎ウイルスのゲノム全領域に対応するDNAである。

#### 【0067】

次に、プラスミドDNA pJFH1及びpJCH1の構造領域と非構造領域の一部を、ネオマイシン耐性遺伝子（neo；ネオマイシンホスホトランスフェラーゼ遺伝子とも称する）及びEMCV-IRES（脳心筋炎ウイルスの内部リボゾーム結合部位）で置換して、プラスミドDNA pSGREP-JFH1及びpSGREP-JCH1をそれぞれ構築した（図1の下段）。この構築手順は、既報（非特許文献7）に従った。具体的には、プラスミドpJFH1及びpJCH1を制限酵素AgeI及びClaIで切断し、その切断部位に、pJFH-1由来の5'NTRからCore領域におよぶ配列とpRSV5NEO由来のネオマイシン耐性遺伝子とをPCR増幅により結合し制限酵素AgeIとPmeIで切断した断片、及びEMCV IRESからNS3領域におよぶ配列をPCR増幅により結合し制限酵素PmeIとClaIで切断した断片を、挿入し連結した。

#### 【0068】

また、pSGREP-JFH1中のNS5B領域について、該領域にコードされるRNAポリ

メラゼの活性中心に相当するアミノ酸モチーフ G D D を G N D に変異させる突然変異を導入して、突然変異プラスミドクローン pSGREP-JFH1/GND を作製した。

#### 【0069】

さらに、pSGREP-JFH1 中の NS5B 領域について、該領域にコードされる R N A ポリメラゼの活性中心に相当するアミノ酸モチーフ G D D を含む連続した 10 アミノ酸配列を欠失させる突然変異を導入して、突然変異プラスミドクローン pSGREP-JFH1/dGDD を作製した。

#### 【0070】

なお上記で作製した突然変異クローン pSGREP-JFH1/GND 及び pSGREP-JFH1/dGDD は、それらにコードされている NS5B タンパク質の活性部位のアミノ酸配列が変異しているため、レプリコン R N A を複製するのに必要な活性 NS5B タンパク質を発現することができない。

#### 【0071】

##### (B) レプリコン R N A の作製

レプリコン R N A 合成に用いる鋳型 D N A を作製するために、上記のとおり構築した発現ベクター pSGREP-JFH1、pSGREP-JCH1、pSGREP-JFH1/GND、pSGREP-JFH1/dGDD を、それぞれ制限酵素 XbaI で切断した。

#### 【0072】

次いで、これらの XbaI 切断断片のそれぞれについて、10~20  $\mu$ g を 50  $\mu$ l の反応液中に含有させ、Mung Bean Nuclease 20 U を用いて 30℃ で 30 分間インキュベートすることにより、さらに処理した。Mung Bean Nuclease は、二本鎖 D N A 中の一本鎖部分を選択的に分解する反応を触媒する酵素である。通常、上記 XbaI 切断断片をそのまま鋳型として用いて R N A 合成を行うと、XbaI の認識配列の一部である CUGA の 4 塩基が 3' 末端に余分に付加されたレプリコン R N A が合成されてしまう。そこで本実施例では、XbaI 切断断片を Mung Bean Nuclease で処理することにより、XbaI 切断断片から CUGA の 4 塩基を除去した。この後、XbaI 切断断片を含む Mung Bean Nuclease 処理後の溶液について、通常法に従ったタンパク質除去処理により、CUGA の 4 塩基が除去された XbaI 切断断片を精製して、これを鋳型 D N A とした。

## 【0073】

次に、この鋳型DNAから、T7 RNAポリメラーゼを用いてRNAをin vitro合成した。このRNA合成にはAmbion社のMEGAscriptを用いた。鋳型DNAを0.5～1.0マイクログラム含む反応液20 $\mu$ lを製造業者の使用説明書に従って反応させた。

## 【0074】

RNA合成終了後、反応溶液にDNase (2 U) を添加して37℃で15分間反応させた後、さらに酸性フェノールによるRNA抽出を行って、鋳型DNAを除去した。このようにしてpSGREP-JFH1、pSGREP-JCH1、pSGREP-JFH1/GND、pSGREP-JFH1/dGDDに由来する上述の鋳型DNAから合成したRNA (レプリコンRNA) を、それぞれrSGREP-JFH1、rSGREP-JCH1、rSGREP-JFH1/GND、rSGREP-JFH1/dGDDと命名した。これらのレプリコンRNAの塩基配列を、rSGREP-JFH1については配列番号1及び図2、rSGREP-JCH1については配列番号2及び図3、rSGREP-JFH1/GNDについては配列番号7、rSGREP-JFH1/dGDDについては配列番号8に示す。

## 【0075】

[実施例2] レプリコン複製細胞クローンの樹立

(C) レプリコンRNAのトランスフェクション、トランスフェクション細胞のコロニー形成能の測定、及び細胞クローンの樹立

上記の合成レプリコンRNA (rSGREP-JFH1、rSGREP-JCH1、rSGREP-JFH1/GND、rSGREP-JFH1/dGDD) のそれぞれについて、様々な量のレプリコンRNAをHuh7細胞から抽出したトータル細胞性RNAと混合して、RNA総量が10 $\mu$ gとなるように調製した。次いでその混合RNAをエレクトロポレーション法によりHuh7細胞に導入した。エレクトロポレーション処理を行ったHuh7細胞を培養ディッシュに播種し、16時間から24時間培養した後に、培養ディッシュにG418 (ネオマイシン) を様々な濃度で添加した。その後、週に2回培養液を交換しながら培養を継続した。播種時から21日間培養した後、クリスタルバイオレットで生存細胞を染色した。染色されたコロニー数を計測し、トランスフェクションしたレプリコンRNA量1 $\mu$ gあたりに得られたコロニー数を計算した。

## 【0076】

コロニー形成が認められたrSGREP-JFH1又はrSGREP-JCH1トランスフェクション細胞については、上記の培養21日後の培養ディッシュからさらに生存細胞のコロニーをクローン化し、培養を継続した。このようなコロニーのクローニングにより、細胞クローンを複数株樹立することができた。

#### 【0077】

樹立した細胞クローンについては、後述の実施例4と同様にして、複製レプリコンRNAの検出、ネオマイシン耐性遺伝子の宿主ゲノムDNAへの組み込みの有無の確認、及びHCVタンパク質の発現の確認を行った。細胞中でのレプリコンの複製が確認された細胞クローンを、レプリコン複製細胞クローンとした。

#### 【0078】

##### (D) 各トランスフェクション細胞におけるコロニー形成能

上記のトランスフェクションの結果、トランスフェクションしたレプリコンRNA 1  $\mu$ g当たりのコロニー形成能は、rSGREP-JFH1をトランスフェクションしたHuh7細胞では、G418濃度が1.0 mg/mlの場合、94700 CFU (Colony Forming Unit; コロニー形成単位)/ $\mu$ g  $\cdot$  RNAであった(図4の左列)。これに対して、rSGREP-JFH1/dGDD、rSGREP-JFH1/GNDをそれぞれトランスフェクションしたHuh7細胞では、コロニー形成が認められなかった(図4の中央列及び右列)。このことは、rSGREP-JFH1レプリコンRNAをトランスフェクションしたHuh7細胞のコロニー形成能は、rSGREP-JFH1から発現されるNS5B (RNAポリメラーゼ)の活性に依存していることを示す。つまり、コロニーを形成した細胞では、rSGREP-JFH1から発現されるNS5BのはたらきによりrSGREP-JFH1レプリコンRNAが自律複製することによって、ネオマイシン耐性遺伝子が持続的に発現されG418耐性が維持される結果、細胞増殖が可能になったものと考えられた。

#### 【0079】

一方、rSGREP-JCH1をトランスフェクションしたHuh7細胞では、G418濃度が1～0.5mg/mlの場合にはコロニー形成が認められなかった(図5)。G418濃度を0.25 mg/mlに下げた場合には、rSGREP-JCH1をトランスフェクションしたHuh7細胞でもコロニー形成が認められた。

#### 【0080】

さらに、上記(B)で得られた発現ベクターpSGREP-JFH1のXbaI切断断片をMung Bean Nucleaseで処理せずにRNA合成の鋳型DNAとして用いてレプリコンRNAを合成し、それを上記(C)と同様にしてHuh7細胞にトランスフェクションした。Mung Bean Nuclease処理を行わずに作製したこのレプリコンRNAには、CU GAの4塩基が3'末端に余分に付加されていた。

#### 【0081】

この結果、Mung Bean Nucleaseで処理せずに作製したレプリコンRNAをトランスフェクションしたHuh7細胞のコロニー形成能は、512 CFU/ $\mu$ g・RNAに低下した(図6の左側)。このことから、レプリコンRNAの3'末端の配列がトランスフェクションした細胞のコロニー形成能に影響を及ぼすことが明らかになった。

#### 【0082】

##### [実施例3]

#### (E) レプリコン複製細胞由来の複製レプリコンRNAの再トランスフェクション

実施例2に従ってrSGREP-JFH1のHuh7細胞へのトランスフェクションにより樹立したレプリコン複製細胞クローンから、常法により全RNA(トータルRNA)を抽出した。この細胞性RNAに含まれる複製レプリコンRNAのコピー数を、ノーザンブロット解析及び定量的RT-PCR法により決定した。

#### 【0083】

ノーザンブロット解析は、Molecular Cloning, A laboratory Manual, 2nd edition, J. Sambrook, E.F. Fritsch, T. Maniatis著、Cold Spring Harbor Laboratory Press (1989)の記載に従って行った。細胞から抽出したRNAを変性アガロース電気泳動に供し、泳動終了後に該RNAをポジティブチャージナイロン膜に転写した。pSGREP-JFH1から作製した<sup>32</sup>PラベルしたDNAまたはRNAプローブを、前記のとおり膜に転写したRNAに対しハイブリダイゼーションさせ、次いでその膜を洗浄し、それをフィルムに感光させることにより、レプリコン特異的なRNAバンドを検出した。

#### 【0084】

レプリコンRNAの定量的RT-PCRによる検出は、Takeuchi T, Katsume A, Tanaka T, Abe A, Inoue K, Tsukiyama-Kohara K, Kawaguchi R, Tanaka S, Kohara M. Real-Time detection system for quantification of Hepatitis C virus genome. Gastroenterology 116: 636-642 (1999) に従いHCV RNAの5'非翻訳領域のRNAを検出することによりおこなった。具体的には、細胞から抽出したRNAに含まれるレプリコンRNAを、合成プライマー、R6-130-S17, 5'-CGGGAGAGC CATAGTGG-3' (配列番号13)、R6-290-R19, 5'-AGTACCACAAGGCCTTTTCG-3' (配列番号14)、TaqMan Probe, R6-148-S21FT, 5'-CTGCGGAACCGGTGAGTACAC-3' (配列番号15) とEZ rTth RNA PCR kitを用いてPCR増幅し、次いでABI Prism 7700 sequence detector systemにより検出した。

#### 【0085】

次に、上記レプリコン複製細胞クローンのうちのクローン6、及びプールクローン (コロニー形成したレプリコン複製細胞を1ディッシュ分集めて培養した細胞) から抽出したトータル細胞性RNAの一部を、再トランスフェクションにより新たなHuh7細胞へ導入した。トランスフェクションに用いたトータル細胞性RNAは、上記で測定したレプリコンRNAのコピー数に基づき、 $1 \times 10^7$ コピーのレプリコンRNAを含むように調製した。トランスフェクションを上記(C)と同様の手法で行い、次いでG418濃度が1mg/mlの条件下で選択培養したところ、レプリコン複製細胞のコロニー形成がみとめられた (図7)。この場合のコロニー形成能は、得られたコロニー数から計算すると、トランスフェクションに用いたレプリコンRNAの $1 \times 10^6$ コピー当たり1コロニー以上であった。

#### 【0086】

一方、pSGREP-JFH1を鋳型としてT7 RNAポリメラーゼにより試験管内で合成したin vitro合成RNAのコピー数は、RNAの重量とRNAの長さから計算すると約 $2 \times 10^{11}$ コピー/ $\mu\text{g} \cdot \text{RNA}$ であり、このin vitro合成RNAを上記と同様にしてトランスフェクションに用いた場合のコロニー形成能は、 $5 \times 10^7$ コピー当たり1コロニーであった。このような結果から、レプリコン複製細胞から抽出した細胞由来RNAとin vitro合成RNAを同じコピー数のレプリコンRNAとしてHuh7細胞にトランスフェクションした場合、Huh7細胞内で複製されたレプ

リコンRNAを用いると、コロニー形成能がin vitro合成RNAと比べて50倍程度高いことが明らかになった。

#### 【0087】

##### [実施例4]

##### (F) レプリコンRNAの検出

上記(E)に従ってrSGREP-JFH1のHuh7細胞へのトランスフェクションにより樹立したレプリコン複製細胞クローンから取得したトータルRNAを新たなHuh7細胞に再トランスフェクションして樹立した細胞クローン[クローン1～11]及びそれらのプールクローン(コロニー形成した細胞クローンを1ディッシュ分集めて培養した細胞)から、酸性フェノール抽出法によりトータルRNAを抽出した。次いでこのトータルRNAをノーザンブロット法により解析した。プローブとしてはpSGREP-JFH1特異的プローブを用いた。対照としては、トランスフェクションを行っていないHuh7細胞から同様に抽出したトータルRNA(図8中、「Huh7」として示す)、Huh7細胞から抽出したトータルRNAに試験管内で合成したレプリコンRNAを10の7乗コピー加えたサンプル(図8中、「 $10^7$ 」として示す)、及びHuh7細胞から抽出したトータルRNAに試験管内で合成したレプリコンRNAを10の8乗コピー加えたサンプル(図8中、「 $10^8$ 」として示す)を用いた。図8中、1～11は細胞クローンの番号である。

#### 【0088】

この結果、rSGREP-JFH1と同程度の大きさのRNAがpSGREP-JFH1特異的プローブにより検出された(図8)。これにより、最初にトランスフェクションしたrSGREP-JFH1に由来するレプリコンRNAが細胞クローン内で複製増殖していることが確認された。また細胞クローン間で、複製レプリコンRNAの量に差があることが示された。図8中、例えば、クローン2、6、9、10は複製レプリコンRNAの量が多く、クローン4、8、11は複製レプリコンRNAの量が少なかった。

#### 【0089】

##### (G) ネオマイシン耐性遺伝子のゲノムDNAへの組み込みの有無の確認

実施例3の手順と同様にして作製した、レプリコンRNAを再トランスフェクションして得られた細胞クローンについて、その細胞クローンのG418に対する耐

性がネオマイシン耐性遺伝子のゲノムへの組み込みによるものでないことを確認するために、ネオマイシン耐性遺伝子特異的プライマー（センスプライマー、NEO-S3: 5'-AACAAGATGGATTGCACGCA-3'（配列番号16）、アンチセンスプライマー、NEO-R: 5'-CGTCAAGAAGGCGATAGAAG-3'（配列番号17））を用いて、細胞クローンから抽出した宿主細胞のゲノムDNAを鋳型とするPCR増幅を行った。用いた細胞クローンは、rSGREP-JFH1由来の複製レプリコンRNAを再トランスフェクションして得られた細胞クローン1～8（rSGREP-JFH1由来細胞クローン1～8）、及びrSGREP-JCH1由来の複製レプリコンRNAを再トランスフェクションして得られた細胞クローン1～6（rSGREP-JCH1由来細胞クローン1～6）であった。この結果、図9に示すとおり、rSGREP-JFH1由来細胞クローンについては、調べた8クローン中、ネオマイシン耐性遺伝子の増幅が示された陽性クローンは認められなかった。rSGREP-JCH1由来細胞クローンについては、調べた6クローン中1クローンのみが陽性であった（図9中、右側写真のレーン3）。この陽性クローンは、rSGREP-JCH1由来の複製レプリコンRNA中のネオマイシン耐性遺伝子が宿主細胞のゲノムDNA中に組み込まれたことにより、G418耐性を獲得したと考えられた。この陽性クローンにおいては、他のクローンとは異なり、レプリコンRNA自体は細胞内で自律複製していないものと考えられた。このことは、次の(H)に示す実験で、この陽性クローンからHCVタンパク質が検出されなかったことにより確認された。

#### 【0090】

##### (H) HCVタンパク質の検出

rSGREP-JFH1及びrSGREP-JCH1トランスフェクション細胞クローンから常法によりタンパク質を抽出して、SDS-PAGE及びウエスタンブロット法により解析した（図10）。調べた細胞クローンは、上記(G)で用いたものと同じであり、rSGREP-JFH1由来細胞クローン1～8、及びrSGREP-JCH1由来細胞クローン1～6である。また、NS3遺伝子を含む発現プラスミドDNAをHuh7細胞にトランジェントにトランスフェクションして得られた細胞抽出液を陽性対照（NS3タンパク質）とした。さらに、トランスフェクションしていないHuh7細胞から抽出したタンパク質を陰性対照として用いた。それぞれの細胞クローンから抽出したタンパク質試



料をPVDF膜 (Immobilon- P , Millipore社製) にブロッティングし、抗NS3特異的抗体 (Dr. Moradpour より分与されたもの; Wolk B, et al, J. Virology. 2000; 74: 2293-2304) を用いて複製レプリコンRNAにコードされているNS3タンパク質を検出した。図10に示されるとおり、rSGREP-JFH1由来細胞クローン1～8及びrSGREP-JCH1由来細胞クローン1、2、4～6では、陽性対照と同じ大きさのタンパク質が検出された。なおrSGREP-JCH1由来細胞クローン3 (上記(G)で陽性クローンとして検出されたクローン) では、NS3タンパク質の発現は検出されなかった。すなわち、rSGREP-JCH1由来細胞クローン3では、レプリコンRNAの複製は確認されなかった。トランスフェクションしていないHuh7細胞でNS3タンパク質が検出されなかったため、NS3タンパク質が検出された細胞クローンでは、トランスフェクションされたレプリコンRNAが自律複製することによりNS3タンパク質が発現されていることが判明した。

#### 【0091】

なお、C型肝炎患者の血清を抗体として用いることにより、上記でNS3タンパク質の発現が確認された各細胞クローンについて、レプリコンRNAからのNS5aタンパク質の発現も確認した。

#### 【0092】

以上の(G)及び(H)の結果から、レプリコンRNAをトランスフェクションして樹立した細胞クローンでは、レプリコンRNAが複製されていることが確認された。

#### 【0093】

##### [実施例5]

##### (I) 適合変異の解析

実施例3に従って、rSGREP-JFH1のHuh7細胞へのトランスフェクションを経て樹立したレプリコン複製細胞クローンから取得したトータルRNAを新たなHuh7細胞に再トランスフェクションして、21の細胞クローンを樹立した。これらの細胞クローンから、常法によりトータルRNAをそれぞれ抽出した。このトータルRNAを鋳型にして逆転写酵素Superscript II (Invitrogen社製) とプライマー9641R-IH (5'-GCACTCTCTGCAGTCATGCGGCTCACGGAC-3' (配列番号18)) によりレ

プリコンRNAに対応するcDNAを合成した。逆転写反応によるcDNA合成のための反応液組成を以下に示す。

## 【0094】

反応液組成	液量( $\mu$ l)
5x 1st strand Buffer	4
2mM dNTP	5
0.1M DTT	1
9651R-IH プライマー (100 $\mu$ M)	1
DW (蒸留水)	6.5
Sample RNA (2 mg/mL)	1
RNasin (Promega社製) (40U/ $\mu$ L)	0.5
<u>Superscript II RT (Invitrogen社製)</u>	<u>1</u>
合計量	20 $\mu$ l

## 【0095】

cDNA合成反応としては、まず上記のRNasinとSuperscript II以外の試薬を混合して最初の反応液を調製し、それを90℃で3分間加熱した後、氷上で冷却した。その後、この反応液に RNasinとSuperscript IIを添加して42℃で1時間反応させた後、さらに70℃で15分間反応させた。

## 【0096】

さらに、このようにして得られたcDNAについて、以下の手順により5組のプライマーセットを用いるPCR増幅を行って、レプリコンRNAのほぼ全領域にわたるDNA増幅断片を得た。用いたプライマーセット及びその各々により増幅される領域を下記の表1及び表2に示す。

## 【0097】

【表 1】

増幅断片の名称	プライマーセット		増幅領域
	プライマー 1	プライマー 2	
A/	42S-IH	433R-neo	41 - 470
B/	C/S17ssp	4680R-IH	28 - 3026
C/	4534S-IH	7279R-IH	2880 - 5625
D/	7198S-IH	9367R-IH	5544 - 7713
E/	9247S-NF	9576R-NF	7597 - 7960

【0098】

なお表 1 中、増幅領域は、rSGREP-JFH1 (配列番号 1) において対応する塩基番号で示した。

【0099】

【表 2】

プライマーの名称	塩基配列 (5' → 3')	配列番号
42S-IH	CCCCTGTGAGGAAGTCTTCTCACGC	配列番号 19
C/S17ssp	CCGGGAGAGCCATAGTGGTCTGCG	配列番号 20
4534S-IH	CCACTCAAAGAAAAAGTGTGACGAGCTCGC	配列番号 21
7198S-IH	GGCTTGGGCACGGCCTGA	配列番号 22
9247S-NF	GCGGTGAAGACCAAGCTCAAACCTCACTCCA	配列番号 23
433R-neo	AGAACCTGCGTGCAATCCATC	配列番号 24
4680R-IH	CCCGTCATGAGGGCGTGGTGGC	配列番号 25
7279R-IH	ACCAGCAACGGTGGGCGGTGGTAATC	配列番号 26
9367R-RI	GGCACGCGACACGCTGTG	配列番号 27
9576R-NF	AGCTAGCCGTGACTAGGGCTAAGATGGAGC	配列番号 28

【0100】

この PCR 反応における反応液組成は以下のとおりである。

【0101】

反応液組成	液量 ( $\mu$ l)
プライマー 1 (10 $\mu$ M)	1.0
プライマー 2 (10 $\mu$ M)	1.0
2.5mM dNTPs	5.0
10x LA Buffer	5.0

MgCl <sub>2</sub> (25mM)	5.0
LA Taq(TAKARA) (5U/ $\mu$ l)	0.3
DW (蒸留水)	30.7
鋳型 cDNA	2.0
合計量	50 $\mu$ l

## 【0102】

また、PCR反応の条件は、以下のとおりであった：95℃で2分間；98℃で10秒間に続き68℃で8分間を35サイクル；72℃で7分間；4℃で保持。

## 【0103】

以上のようにして得られた各PCR産物の塩基配列を決定し、そのDNA配列対応するRNA配列とrSGREP-JFH1の配列との比較を行った。その結果を表3に示す。

## 【0104】

【表3】

領域	同義的 置換	非同義 置換	全変異
NS3	0	5	5
NS4A	0	2	2
NS4B	0	3	3
NS5A	0	7	7
NS5B	3	5	8
合計	3	22	25

## 【0105】

表3のとおり、21の細胞クローンにおいて認められた塩基変異は全部で25個あったが、そのうち22個はアミノ酸の変異を引き起こす非同義置換であった。これらの変異の種類を、表4に示す。またこれらの変異の非構造領域における

位置を図 11 に示す。

【0106】

【表 4】

クローン名	変異部位			
	塩基番号	塩基変異	アミノ酸変異	アミノ酸番号
C1	7098	A ⇄ G	なし	
	7157	A ⇄ G	Y ⇄ C	2824
C2	4955	C ⇄ U	A ⇄ V	2090
C3	4936	A ⇄ G	T ⇄ A	2084
	5000	A ⇄ G	Y ⇄ C	2105
	7287	A ⇄ G	なし	
	7288	A ⇄ G	M ⇄ V	2868
C4	5901	G ⇄ U	E ⇄ D	2405
	6113	A ⇄ U	H ⇄ L	2476
C5	2890	A ⇄ G	K ⇄ E	1402
C6	7209	A ⇄ G	なし	

【0107】

表 4 及び図 11 において、「C1～C6」は、変異が見出されたレプリコン RNA を有するレプリコン複製細胞クローン C1～C6 を示す。「塩基番号」はレプリコン RNA rSGREP-JFH1 の塩基配列（配列番号 1）中の対応する塩基の番号として示している。「アミノ酸番号」は、JFH-1 クローンにコードされるアミノ酸配列（配列番号 4）中の対応するアミノ酸の番号として示している。変異部位の塩基及びアミノ酸の種類は、通常の表記法に従って記載されている。表 4 に示されるとおり、クローン C2 では、レプリコン RNA 上の配列番号 1 の塩基番号 4955 に相当する塩基が C（シトシン）から U（ウラシル）へ変異し、その結果、配列番号 4 のアミノ酸番号 2090 に相当するアミノ酸が A（アラニン）から V（バリン）へ変異した。

【0108】

また図11に示した変異の位置は、表4に示す塩基番号を付記した縦線で示している。太い縦線は非同義置換、細い縦線は同義置換を表す。

#### 【0109】

なお、アミノ酸の変異を引き起こす塩基変異を全く有しないクローンも2クローンあった。この2クローンについてノーザンブロット解析を行ったところ、これらの2クローンにおいては、レプリコンRNAの複製量が、アミノ酸変異を生じる塩基変異を有するレプリコンRNAを複製している細胞クローンよりも少ないことが示された。つまり、レプリコンRNAにおけるアミノ酸変異を引き起こす塩基変異は、Huh7細胞においてレプリコンRNAの複製効率を増大させる適合変異であると考えられた。

#### 【0110】

##### 【発明の効果】

本発明により、遺伝子型2aのHCV株に由来するHCV-RNAレプリコンが初めて得られた。本発明に係るレプリコン複製細胞は、遺伝子型2aのHCV由来のRNA及びHCVタンパク質を持続的に産生させるための培養系として用いることができる。さらに本発明に係るレプリコン複製細胞は、HCVの複製及び／又はHCVタンパク質の翻訳に影響を及ぼす各種物質をスクリーニングするための試験系として有用である。

#### 【0111】

##### 【配列表】

#### SEQUENCE LISTING

<110> Toray Industries Inc.

Tokyo Metropolitan Organization for Medical Research

Johannes Gutenberg-Universitaet Mainz

<120> Establishment of the genotype 2a Hepatitis C virus subgenomic replicon

<130> P03-0429

<140>

<141>

<160> 28

<170> PatentIn Ver. 2.1

<210> 1

<211> 8024

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: replicon

<220>

<223> Inventor: Wakita, Takaji

Inventor: Kato, Takanobu

Inventor: Date, Tomoko

<400> 1

accugccccc aauaggggacg acacuccgcc augaauacac cccugugag gaacuacugu 60  
cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120  
ccccucccg ggagagccau aguggucugc ggaaccggug aguacaccgg aaaugccggg 180  
aagacugggu ccuuucuugg auaaaccac ucuaugcccg gccauuuggg cgugcccccg 240  
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300  
cgcuugcgag ugccccggga ggucucguag accgugcacc augagcaca auccuaaacc 360  
ucaaagaaaa accaaaagaa acaccaaccg ucgcccgaug auugaacaag auggauugca 420

cgagguucu ccggccgcuu ggguggagag gcuauucggc uaugacuggg cacaacagac 480  
aaucggcugc ucugaugccg ccguguuccg gcugucagcg cagggggcgcc cgguucuuuu 540  
ugucaagacc gaccuguccg gugcccugaa ugaacugcag gacgaggcag cgcggcuauc 600  
guggcuggcc acgacgggcg uuccuugcgc agcugugcuc gacguuguca cugaagcggg 660  
aaggggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720  
uccugccgag aaaguaacca ucauggcuga ugcaaugcgg cggcugcaua cgcuugauc 780  
ggcuaccugc ccuucgacc accaagcgaa acaucgcauc gagcgagcac guacucggau 840  
ggaagccggu cuugucgauc aggaugaucu ggacgaagag caucaggggc ucgcgccagc 900  
cgaacuguuc gccaggcuca aggcgcgcau gcccgcggc gaggaucucg ucgugacca 960  
uggcgaugcc ugcuugccga auaucauggu ggaaauggc cgcuuuucug gauucaucga 1020  
cuguggcccg cugggugugg cggaccgcu ucaggacaua gcguuggcu cccgugauau 1080  
ugcugaagag cuuggcggcg aaugggcuga ccgcuuccuc gugcuuacg guaucgccgc 1140  
ucccgauucg cagcgcaucg ccuucuaucg ccuucugac gaguucuucu gaguuuaaac 1200  
ccucuccuc ccccccccu aacguuacug gccgaagccg cuuggaauaa ggccggugug 1260  
cguuugucua uauguuuuu uccaccauau ugccgucuuu uggcaaugug agggcccga 1320  
aaccuggccc ugucuucug acgagcauuc cuaggggucu uccccucuc gccaaaggaa 1380  
ugcaaggucu guugaauguc gugaaggaag caguuccucu ggaagcuucu ugaagacaaa 1440  
caacgucugu agcgaccuu ugcaggcagc ggaaccccc accuggcgac agguiggucu 1500  
gcggccaaaa gccacgugua uaagauacac cugcaaaggc ggcacaacc cagugccacg 1560  
uugugaguug gauaguugug gaaagaguca aauggcucuc cucaagcgu uucaacaagg 1620  
ggcugaagga ugcccagaag guaccccauu gnaugggauc ugaucugggg ccucggugca 1680  
caugcuuuac auguguuuag ucgagguuaa aaaaacgucu agggcccccg aaccacgggg 1740  
acgugguuuu ccuuugaaaa acacgaugau accauggcuc ccaucacugc uuaugcccag 1800  
caaacacgag gccuccuggg cgccauagug gugaguauga cggggcguga caggacagaa 1860  
caggccgggg aaguccaaau ccuguccaca gucucucagu ccuuccucgg aacaaccauc 1920  
ucggggguuu uguggacugu uuaccacgga gcuggcaaca agacucuagc cggcuuacgg 1980  
gguccgguca cgcagaugua cucgagugcu gagggggacu ugguaggcug gccagcccc 2040  
ccugggacca agucuuugga gccgugcaag uguggagccg ucgaccuaua ucuggucacg 2100  
cggaacgcug augucaucc ggucgggaga cgcggggaga agcggggagc auugcucucc 2160



ccgagaccca uuucgaccuu gaaggggucc ucgggggggc cggugcucug cccuaggggc 2220  
cacgucguug ggcucuuccg agcagcugug ugcucucggg gcguggccaa auccaucgau 2280  
uucauccccg uugagacacu cgacguuguu acaaggucuc ccacuucag ugacaacagc 2340  
acgccaccgg cugugcccca gaccuaucag gucggguacu ugcaugcucc aacuggcagu 2400  
ggaaagagca ccaagguccc ugucgcguau gccgcccagg gguacaaagu acuagugcuu 2460  
aaccucucgg uagcugccac ccugggguuu ggggcguacc uauccaaggc acauggcauc 2520  
aaucccaaca uuaggacugg agucaggacc gugaugaccg gggaggccau cacguacucc 2580  
acauauggca aaauucucgc cgaugggggc ugcgcuaagc gcgccuauga caucaucaua 2640  
ugcgaugaau gccacgcugu ggaugcuacc uccauucucg gcaucggaac gguccuugau 2700  
caagcagaga cagccggggg cagacuaacu gugcuggcua cggccacacc ccccggguca 2760  
gugacaaccc cccaucccga uauagaagag guaggccucg ggcgggaggg ugagaucucc 2820  
uucuauggga gggcgauucc ccuauccugc aucaaggag ggagacaccu gauuuucugc 2880  
cacucaaaga aaaaguguga cgagcucgcg gcggcccuuc ggggcauggg cuugaaugcc 2940  
guggcauacu auagaggguu ggacgucucc auauaccag cucagggaga ugugguguc 3000  
gucgccaccg acgccucau gacgggguac acuggagacu uugacuccgu gaucgacugc 3060  
aanguagcgg ucaccaagc ugucgacuuc agccuggacc ccaccuucac uauaaccaca 3120  
cagacugucc cacaagacgc ugucucacgc agucagcgcc gcgggcgcac agguagagga 3180  
agacagggca cuuauaggua uguuuccacu ggugaacgag ccucaggaau guuugacagu 3240  
guagugcuu gugagugcua cgacgcaggg gcugcguggu acgaucucac accagcggag 3300  
accaccguca ggcuuagagc guauuucac acgcccggcc uaccgugug ucaagaccu 3360  
cuugaauuuu gggaggcagu uuucaccggc cucacacaca uagacgcca cuuccucucc 3420  
caaacaagc aagcggggga gaacuucgcg uaccuaguag ccuaccaagc uacgguguc 3480  
gccagagcca agggccucc cccguccugg gacgcaugu ggaagugccu ggcccaguc 3540  
aagccuacgc uugcggggc cacaccucuc cuguaccguu ugggcccua uaccaaugag 3600  
gucaccuca cacaccugg gacgaaguac aucgccacu gcaugcaagc ugaccuugag 3660  
gucaugacca gcacgugggu ccuagcugga ggaguccugg cagccgucgc cgcauauugc 3720  
cuggcgacug gaugcguuuc caucaucggc cgcuugcacg ucaaccagcg agucgucguu 3780  
gcgcccgaug aggagguccu guaugaggcu uuugaugaga uggaggaaug cgccucuagg 3840  
gcggcucuca ucgaagaggg gcagcggaua gccgagaugu ugaaguccaa gauccaaggc 3900

uugcugcagc aggccucuaa gcaggcccag gacauacaac ccgcuaugca ggcuucaugg 3960  
cccaaagugg aacaauuuug ggccagacac auguggaacu ucauuagcgg cauccaauac 4020  
cucgcaggau ugucaacacu gccagggaac cccgcggugg cuuccaugau ggcauucagu 4080  
gccgcccuca ccaguccguu gucgaccagu accaccaucc uucuacaau caugggaggc 4140  
ugguuagcgu ccagaucgc accaccgcg ggggccaccg gcuuugucgu caguggccug 4200  
gugggggug ccgugggcag cauaggccug gguaaggugc ugguggacau ccuggcagga 4260  
uauggugcgg gcauuucggg ggcccucguc gcauuaaga ucaugucugg cgagaagccc 4320  
ucauggaag augucauaa ucuacugccu gggauccugu cuccgggagc ccugguggug 4380  
ggggucaucu gcgcggccau ucugcgccg cacgugggac cgggggaggg cgcgguccaa 4440  
uggaugaaca ggcuuauugc cuuugcuucc agaggaaacc acgucgccc uacucacuac 4500  
gugacggagu cggaugcguc gcagcgugug acccaacuac uuggcucucu uacuauaacc 4560  
agccuacua gaagacucca caauuggaua acugaggacu gcccacucc augcuccgga 4620  
uccuggcucc gcgacgugug ggacuggguu ugcaccaucu ugacagacuu caaaaauugg 4680  
cugaccucua aauguuucc caagcugccc ggccucccu ucaucucuug ucaaaagggg 4740  
uacaagggug ugugggccgg cacuggcauc augaccacgc gcugcccuug cggcgccaac 4800  
aucucuggca auguccgccu gggcucuaug aggaucacag ggccuaaaac cugcaugaac 4860  
accuggcagg ggaccuuucc uaucaauugc uacacggagg gccagugcg gccgaaacc 4920  
cccacgaacu acaagaccgc caucuggagg guggcggccu cggaguacgc ggaggugacg 4980  
cagcaugggu cguacuccua uguaacagga cugaccacug acaaucugaa aaauccuugc 5040  
caacuaccuu cuccagaguu uuucuccugg guggacggug ugcagaucca uagguuugca 5100  
cccacaccaa agccguuuuu ccgggaugag gucucguucu gcguugggcu uaauuccuau 5160  
gcugucgggu ccagcuucc cugugaaccu gagcccgacg cagacguauu gagguccaug 5220  
cuaacagauc cgccccacau cacggcgagg acugcgcg gcgcguuggc acggggauca 5280  
ccuccaucug aggcgagcuc cucagugagc cagcuaucag caccgucgu gcggggcacc 5340  
ugcaccaccc acagcaacac cuaugacgug gacauggugc augccaaccu gcucauggag 5400  
ggcggugugg cucagacaga gccugagucc agggugcccg uucuggacuu ucucgagcca 5460  
auggccgagg aagagagcga ccuugagccc ucaauaccu cggagugcau gcuccccagg 5520  
agcggguuuc cacgggccuu accggcuugg gcacggccug acuacaacc gccgcucgug 5580  
gaaucgugga ggaggccaga uuaccaaccg cccaccguug cugguugugc ucucggggg 5640

cccaagaagg ccccgacgcc ucccccaagg agacgccgga cagugggucu gagcgagagc 5700  
accauaucag aagcccucca gcaacuggcc aucaagaccu uuggccagcc ccccucgagc 5760  
ggugaugcag gcucguccac gggggcgggc gccgccgaau ccggcggucc gacgucccu 5820  
ggugagccgg ccccucaga gacagguucc gccuccucua ugccccccu cgaggggggag 5880  
ccuggagauc cggaccugga gucugaucag guagagcuuc aaccucuccc ccaggggggg 5940  
gggguagcuc ccgguucggg cucggggucu uggucuacuu gcuccgagga ggacgauacc 6000  
accgugugcu gcuccauguc auacuccugg accggggcuc uaauaacucc cuguagcccc 6060  
gaagaggaaa aguugccaau caaccuuug aguaacucgc uguugcgaua ccuaaacaag 6120  
guguacugua caacaucaaa gagcgccuca cagagggcua aaaagguaac uuuugacagg 6180  
acgcaagugc ucgacgcca uuaugacuca gucuuaaagg acaucaagcu agcggcuucc 6240  
aaggucagcg caaggcuccu caccuuggag gaggcgugcc aguugacucc accccaauuc 6300  
gcaagaucca aguauggauu cggggccaag gagguccgca gcuuguccgg gagggccguu 6360  
aaccacauca aguccgugug gaaggaccuc cuggaagacc cacaacacc aaaucccaca 6420  
accaucaugg caaaaauga gguguucugc guggaccccg ccaagggggg uaagaaacca 6480  
gcucgccuca ucguuuaccc ugaccucggc guccgggucu gcgagaaaau ggcccucua 6540  
gacauuacac aaaagcuucc ucaggcgguu augggagcuu ccuauggcuu ccaguacucc 6600  
ccugcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagga cccaugggu 6660  
uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720  
gaguccauau accaggccug cuccugccc gaggaggccc gcacugccau acacucgcug 6780  
acugagagac uuuacguagg agggcccaug uucaacagca aggguaaac cugcgguuac 6840  
agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 6900  
gugaaagccc uagcggccug caaggcugcg gggauaguug cgcccacaau gcugguau 6960  
ggcgaugacc uaguagucau cucagaaagc caggggacug aggaggacga gcggaaccug 7020  
agagccuua cggaggccau gaccagguac ucugcccuc cuggugauc cccagaccg 7080  
gaauaugacc uggagcuau acauccugu uccucaaau ugucuguggc guugggccc 7140  
cggggccgcc gcagauacua ccugaccaga gaccacaac cuccacucgc ccgggcugcc 7200  
ugggaaacag uuagacacuc ccuaucaau ucauggcugg gaaacauc auccaguauc 7260  
ccaaccauau ggguucgcau gguccuaau acacacuuc ucuccauuc caugguccaa 7320  
gacaccugc accagaaccu caacuuugag auguauggau caguauacuc cgugaauc 7380

uuggaccuuc cagccauaau ugagagguua cacgggcuug acgccuuuuc uaugcacaca 7440  
uacucucacc acgaacugac gcggguggcu ucagcccuca gaaaacuugg ggcgccaccc 7500  
cucagggugu ggaagagucg ggcucgcgca gucagggcgu cccucaucuc ccguggaggg 7560  
aaagcggccg uuugcggccg auaucucuuc aaugggcg ugaagaccaa gcucaaacuc 7620  
acuccauugc cggaggcgcg ccuacuggac uuauccaguu gguuaccgu cggcgccggc 7680  
gggggcgaca uuuuucacag cgugucgcgc gcccgacccc gcucauuacu cuucggccua 7740  
cuccuacuuu ucguaggggu aggccucuuc cuacucggcg cucgguagag cggcacacac 7800  
uagguacacu ccuagcuuaa cuguuccuuu uuuuuuuuuu uuuuuuuuuu uuuuuuuuuu 7860  
uuuuuuuuuu cuuuuuuuuu uuuuuuccuc uuucuuuccu ucucaucuua uucuaucuuc 7920  
uuucuuuggg gcuccaucuu agcccuaguc acggcuagcu gugaaagguc cgugagccgc 7980  
augacugcag agagugccgu aacuggucuc ucugcagauc augu 8024

<210> 2

<211> 8024

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: replicon

<400> 2

acccgccccu aaugggggcg acacuccgcc augaaucacu cccugugag gaacuacugu 60  
cuucacgcag aaagcgucua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120  
ccccucccg ggagagccau aguggucugc ggaaccggug aguacaccgg aaugccggg 180  
aagacugggu ccuuucuuug auaaaccac ucaugcccg gccauuuggg cgugcccccg 240  
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300  
ugcuugcgag ugccccggga ggucucguag accgugcacc augagcaca aucccaaacc 360  
ucaaaagaaa accaaaagaa acacuaaccg ucgcccuaug auugaacaag auggauugca 420

cgcagguucu ccggccgcuu ggguggagag gcuauucggc uaugacuggg cacaacagac 480  
aaucggcugc ucugaugccg ccguguuccg gcugucagcg caggggccc cgguucuuu 540  
ugucaagacc gaccuguccg gugcccugaa ugaacugcag gacgaggcag cgcggcuauc 600  
guggcuggcc acgacgggcg uuccuugcgc agcugugcuc gacguugua cugaagcggg 660  
aaggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720  
uccugccgag aaaguaacca ucauggcuga ugcaaugcgg cggcugcaua cgcuugauc 780  
ggcuaccugc ccuucgacc accaagcgaa acaucgcau gagcgagcac guacucggau 840  
ggaagccggu cuugucgauc aggaugaucu ggacgaagag caucaggggc ucgcgccagc 900  
cgaacuguuc gccaggcuca aggcgcgcau gcccgcggc gaggauucg ucgugacca 960  
uggcgaugcc ugcugccga auaucauggu ggaaauggc cgcuuucug gauucaucga 1020  
cuguggccgg cugggugugg cggaccgcu ucaggacaua gcguuggcu cccgugauau 1080  
ugcugaagag cuuggcggcg aaugggcuga ccgcuuccuc gugcuuacg guaucgccgc 1140  
ucccgauucg cagcgcaucg ccuucuaucg ccuucugac gaguucucu gaguuuaac 1200  
ccucuccuc cccccccu aacguuacug gccgaagccg cuuggaauaa ggccggugug 1260  
cguuugucia uauguauuu uccaccauau ugccgucuuu uggcaaugug agggcccga 1320  
aaccuggccc ugucuucug acgagcauuc cuaggggucu uccccucuc gccaaaggaa 1380  
ugcaaggucu guugaaugc gugaaggaag caguuccucu ggaagcuuc ugaagacaaa 1440  
caacgucugu agcgaccuu ugcaggcagc ggaaccccc accuggcgac aggugccuc 1500  
gcggccaaaa gccacgugua uaagauacac cugcaaaggc ggcacaacc cagugccacg 1560  
uugugaguug gauaguug gaaagaguc aauggcucuc cucaagcgu uucaacaagg 1620  
ggcugaagga ugcccagaag guaccccau guaugggauc ugaucugggg ccucggugca 1680  
caugcuuac auguguuag ucgagguua aaaaacguc agggcccccg aaccacgggg 1740  
acgugguuu ccuuugaaaa acacgauau accauggcc ccaucaccgc uuacgccag 1800  
cagacacgag gucucuuggg cucuauagug gugagcaua cggggcguga caagacagaa 1860  
caggccgggg agguccaagu ccugucca gucacucagu ccuuccucg acauccauu 1920  
ucgggggucu uauggacugu uuaccacgga gcuggcaaca agacacuagc cggcucgcgg 1980  
ggcccggua cgcagaugua cucgagcgcc gagggggacu uggucgggug gccagcccu 2040  
ccugggacca aaucuuugga gccguguacg uguggagcgg ucgaccugua uuuggucacg 2100  
cggaacgcug augucaucc ggcucgaaga cgcggggaca agcggggagc gcugcucucc 2160

ccgagacccc uuucgaccuu gaagggggucc ucgggggggac cugugcuuug cccuaggggc 2220  
cacgcugucg gaauuuccg ggcagcugug ugcucucggg guguggcuaa guccauagau 2280  
uucauccccg uugagacgcu cgacaucguc acgcggucuc ccaccuuuag ugacaacagc 2340  
acaccaccag cugugcccca gaccuauacg guggggguacu ugcacgcccc cacuggcagu 2400  
ggaaaaagca ccaagguccc cgucgcguac gccgcccagg gguauaaagu gcuggugcuc 2460  
aaucccucgg uggcugccac ccugggauuu ggggcguacu uguccaaggc acauggcauc 2520  
aaccccaaca uuaggacugg agucagaacu gugacgaccg gggagcccau uacauacucc 2580  
acguauggua aaauccucgc cgaugggggc ugcgcaggcg gcgccuauga caucaucaua 2640  
ugcgaugaau gccacucugu ggaugcuacc acuauucucg gcaucgggac aguccuugac 2700  
caagcagaga cagccggggg caggcuaacu guacuggcca cggccacgcc ccccgggucg 2760  
gugacaaccc cccaucccaa uauagaggag guagcccucg gacaggaggg ugagaucucc 2820  
uucuauggga gggcguuucc ccugucuuac aucaaggag ggaggcacuu gauuuucugc 2880  
cacucaaaga aaaaguguga cgagcucgca acggcccuuc ggggcauggg cuugaacgcu 2940  
guggcauauu acagaggguu ggacgucucc auauuaccaa cucaaggaga uguggugguc 3000  
guugccaccg acgccucau gacggggauu acuggagacu uugacuccgu gaucgacugc 3060  
aacguagcgg ucaccaggc cguagacuuc agccuggacc ccaccuucac uauaaccaca 3120  
cagacugucc cgcaagacgc ugucucacgu agucagcgcc gagggcgcac ggguagagga 3180  
agacugggca uuuauaggua uguuuccacu ggugagcgag ccucaggaau guuugacagu 3240  
guaguacucu gugagugcua cgacgcagga gcugcuuggu augagcucuc accaguggag 3300  
acgaccguca ggcucagggc guauuucaac acgccuggcu ugccugugug ccaggaccac 3360  
cuugaguuuu gggaggcagu uuucaccggc cucacacaca uagacgcua uuuccuuucc 3420  
cagacaaagc agucggggga aaauuucgca uacuuaguag ccuaucaggc cacagugugc 3480  
gccagggcca aagcgcccc cccguccugg gacgucaugu ggaagugcuu gacucgacuc 3540  
aagcccacgc uuuggggccc uacaccucuc cuguaccguu ugggcucugu uaccaacgag 3600  
gucaccuua cacaccgcu gacaaaauac aucgccacau gcaugcaagc ugaccucgag 3660  
gucaugacca gcacgugggu ccuggcuggg ggagucuuag cagccgucgc cgcguauugc 3720  
uuagcgaccg gguguguuuc caucauuggc cguuuacaca ucaaccagcg agcugucguc 3780  
gcuccggaca aggagguccu cuaugaggcu uuugaugaga uggaggaaug ugccuccaga 3840  
gcggcucucc uugaagaggg gcagcggaua gccgagaugc ugaaguccaa gauccaaggc 3900

uuauugcagc aagccucuaa acaggcccag gacauacaac ccgcugugca agcuucgugg 3960  
cccaagaugg agcaauucug ggccaaacau auguggaacu ucauaagcgg cauucaguac 4020  
cucgcaggac ugucaacacu gccagggaac ccugcugugg cuuccaugau ggcauucagc 4080  
gccgcccuca ccaguccguu gucaacuagc accaccaucc uucuaacau ucuggggggc 4140  
uggcuggcgu cccaaauugc gccacccgcg gggggccacug gcuuuguugu caguggccug 4200  
gugggagcug cuguuggcag cauaggcuug gguaaagugc ugguggacau ccuggcaggg 4260  
uauggugcgg gcauuucggg ggcccucguc gcguuaaga ucaugucugg cgagaagccc 4320  
uccauggagg augucauaa cuugcugccu gggauucugu cuccaggugc ucugguggug 4380  
ggagucaucu gcgcggccau ucugcgccgc caugugggac cgggggaagg cgcgguccaa 4440  
uggaugaaca ggcuuauugc cuucgcuucc agaggaaacc acgucgccc uacucacuac 4500  
gugacggagu cggaugcguc gcagcguguc acccaacugc uggcucucu cacuaauacu 4560  
agucuacuca ggagacuua caacuggauc acugaggauu gcccacucc augcgccggc 4620  
ucguggcucc gcgaugugug ggacuggguc uguaccaucc uaacagacuu uaagaacugg 4680  
cugaccucca agcuguucc aaagaugccu ggccucccu uuauucuuug ccaaagggg 4740  
uacaagggcg ugugggccgg cacuggcauc augaccacac gaugcccug cggcgccaac 4800  
aucucuggca acguccgcuu gggcucuaug agaauacag gacccaaaac cugcaugaac 4860  
accuggcagg ggaccuuucc uaucaauugu uauacagaag gccagugcuu gccgaaacc 4920  
gcguuaaacu ucaagaccgc caucuggaga guggcgccu cagaguacgc ggaagugacg 4980  
cagcacggau cauaugccua uauaacaggg cugaccacug acaacuuaaa agucccuugc 5040  
caacucccu cuccagagu uuucucuugg guggacggag uacaaaucca uagguccgcc 5100  
cccacacca agccguuuu ccgggaugag gucucguuca gcguugggcu caauucauuu 5160  
gucgucgggu cucagcuucc cugugaccu gagcccgaca cugagguagu gauguccaug 5220  
cuaacagacc cauccauau cacggcggag gcugcagcgc ggcguuagc gcggggguca 5280  
ccccaucug aggcaagcuc cucagcgagc cagcugucgg cgccaucgu gcgagccacc 5340  
ugcaccacc acgguaggac cuaugaugug gacauggugg augccaaccu guucaugggg 5400  
ggcggcguga uucggauaga gucugagucc aaaguggucg uucuggacuc ccucgacuca 5460  
augaccgagg aagagggcga ccuugagccu ucaguaccu cggaguauau gcuccccagg 5520  
aagagguucc caccggccu accggcuugg gcgcggccug auuacaacc accgcuugug 5580  
gaucgugga agaggccaga uuaccaacca cccacuguug cgggcugugc ucucuccccc 5640

cccaaaaaga ccccgacgcc uccuccaagg agacgccgga cagugggucu gagcgagagc 5700  
accauaggag augcccuca acagcuggcc aucaaguucu uuggccagcc cccccaagc 5760  
ggcgauucag gccuuuccac gggggcggac gccgccgacu ccggcgaucg gacaccccu 5820  
gacgaguugg cucuuucgga gacagguucu accuccucca ugccccccu cgagggggag 5880  
ccuggggacc cagaccugga gccugagcag guagagcuuc aaccuccucc ccaggggggg 5940  
gaggcagcuc ccggcucgga cucggggucc uggucuacuu gcuccgagga ggaugacucc 6000  
gucgugugcu gcuccauguc auauuccugg accggggcuc uaauaacucc uuguagcccc 6060  
gaagaggaaa aguugccaau uaacuccuug agcaacucgc uguugcgaua ccuaaacaag 6120  
guauacugua cuacaucaaa gagugccuca cuaagggcua aaaagguaac uuuugauagg 6180  
augcaagugc ucgacgccua uuaugauuca gucuuaaagg acaucaagcu agcggccucc 6240  
aaggucagcg caaggcuccu caccuuagag gaggcgugcc aaugacccc accccacucu 6300  
gcaagaacca aguauggguu uggggcuaag gagguccgca gcuuguccgg gagggccguc 6360  
aaccacauca aguccgugug gaaggaccuc uuggaagacu cacaacacc aauccuaca 6420  
accaucaugg caaaaaaga gguguucugc guggacccc ccaagggggg uaaaaacca 6480  
gcucgccuua ucguuuacc ucgaccucggc gucagggucu gcgagaagau ggccuuuau 6540  
gaugucacac aaaagcuucc ucaggcggug augggggcuu cuuauggcuu ccaguacucc 6600  
cccgucagc ggguggaguu ucucuugaag gcaugggcgg aaaagagaga ccuaugggu 6660  
uuuucguaug auaccggaug cuuugacuca accgucacug agagagacau caggacugag 6720  
gaguccauau accaggccug cuccuuacc gaggaggccc gaacugccau acacucgcug 6780  
acugagagac ucuauuggg agggcccaug uucaacagca agggccaguc cugcgguuac 6840  
aggcguugcc gcgccagcgg ggugcuuacc acuaguaug ggaacaccu cacaugcuau 6900  
guaaaagccc uagcggcuug caaggcugcg gggauaaug cgccacgau gcugguau 6960  
ggcgacgacu uggucguc cuagaaaagc caggggacug aggaggacga gcggaaccug 7020  
agagccuua cggaggcuau gaccaggau ucugcccuc cuggugaccc cccagaccg 7080  
gaauaugacc uggagcuau aacauuugu uccucaaagc ugucuguggc acuuggccca 7140  
cagggccgcc gcagauacua ccugaccaga gacccacca cuucaauugc ccgggcugcc 7200  
ugggaaacag uuagacacuc ccugucau ucauggcugg gaaacaucau ccaguacgu 7260  
ccaaccuau ggguucgcu gguccugaug acacacuuc ucuccauuc cauggcccag 7320  
gacaccuag accagaaccu uaacuugaa auguacggau cgguuacuc cgugaguccu 7380



cuggaccucc cagccauaau ugaaagguua cacgggcuug acgccuucuc ucugcacaca 7440  
uacacucccc acgaacugac gcggguggcu ucagcccua gaaaacuugg ggcgccaccc 7500  
cucagagcgu ggaagagucg ggcgcgugca guuagggcgu ccucaucuc ccgugggggg 7560  
agggcggccg uuugcgucg guaccucuuc aacugggcgg ugaagaccaa gcucaaacuc 7620  
acuccuuugc cggaggcacg ccuccuggau uuguccagu gguuuaccgu cggcgccggc 7680  
gggggcgaca uuuaucacag cgugucgcu gcccacccc gccuauuacu ccuagccua 7740  
cuccuacuuu cuguaggggu aggccucuuc cuacucccg cucgauagag cggcacacau 7800  
uagcuacacu ccuagcuua cuguuccuuu uuuuuuuuuu uuuuuuuuuu uuuuuuuuuu 7860  
uuuuuuuuu cuuuuuuuu uuuuuccuc uuucuuccu ucucaucua uucuacuuu 7920  
uuucuuggug gcuccaucu agcccuaguc acggcuagcu gugaaagguc cgugagccgc 7980  
augacugcag agagugccgu aacuggucuc ucugcagau augu 8024

<210> 3

<211> 9678

<212> DNA

<213> Hepatitis C virus

<220>

<221> CDS

<222> (341)..(9442)

<400> 3

acctgccct aatagggcg acactccgc atgaatcact ccctgtgag gaactactgt 60

cttcacgcag aaagcgcta gccatggcgt tagtatgagt gtcgtacag ctccaggccc 120

cccctcccg ggagagccat agtggctgc ggaaccggtg agtacaccgg aattgccggg 180

aagactgggt cctttcttgg ataaaccac tctatgcccg gccatttggg cgtgcccccg 240

caagactgct agccgagtag cgttgggttg cgaaaggcct tgttgtactg cctgataggg 300

cgcttgcgag tgccccggga ggtctcgtag accgtgcacc atg agc aca aat cct 355

Met Ser Thr Asn Pro

1 5

aaa cct caa aga aaa acc aaa aga aac acc aac cgt cgc cca gaa gac 403

Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Glu Asp

10 15 20

gtt aag ttc ccg ggc ggc ggc cag atc gtt ggc gga gta tac ttg ttg 451

Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Gly Val Tyr Leu Leu

25 30 35

ccg cgc agg ggc ccc agg ttg ggt gtg cgc acg aca agg aaa act tcg 499

Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Thr Thr Arg Lys Thr Ser

40 45 50

gag cgg tcc cag cca cgt ggg aga cgc cag ccc atc ccc aaa gat cgg 547

Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro Ile Pro Lys Asp Arg

55 60 65

cgc tcc act ggc aag gcc tgg gga aaa cca ggt cgc ccc tgg ccc cta 595

Arg Ser Thr Gly Lys Ala Trp Gly Lys Pro Gly Arg Pro Trp Pro Leu

70 75 80 85

tat ggg aat gag gga ctc ggc tgg gca gga tgg ctc ctg tcc ccc cga 643

Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp Leu Leu Ser Pro Arg

90

95

100

ggc tct cgc ccc tcc tgg ggc ccc act gac ccc cgg cat agg tcg cgc 691

Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro Arg His Arg Ser Arg

105

110

115

aac gtg ggt aaa gtc atc gac acc cta acg tgt ggc ttt gcc gac ctc 739

Asn Val Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu

120

125

130

atg ggg tac atc ccc gtc gta ggc gcc ccg ctt agt ggc gcc gcc aga 787

Met Gly Tyr Ile Pro Val Val Gly Ala Pro Leu Ser Gly Ala Ala Arg

135

140

145

gct gtc gcg cac ggc gtg aga gtc ctg gag gac ggg gtt aat tat gca 835

Ala Val Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala

150

155

160

165

aca ggg aac cta ccc ggt ttc ccc ttt tct atc ttc ttg ctg gcc ctg 883

Thr Gly Asn Leu Pro Gly Phe Pro Phe Ser Ile Phe Leu Leu Ala Leu

170

175

180

ttg tcc tgc atc acc gtt ccg gtc tct gct gcc cag gtg aag aat acc 931

Leu Ser Cys Ile Thr Val Pro Val Ser Ala Ala Gln Val Lys Asn Thr

185

190

195

agt agc agc tac atg gtg acc aat gac tgc tcc aat gac agc atc act 979

Ser Ser Ser Tyr Met Val Thr Asn Asp Cys Ser Asn Asp Ser Ile Thr

200	205	210	
tgg cag ctc gag gct gcg gtt ctc cac gtc ccc ggg tgc gtc ccg tgc			1027
Trp Gln Leu Glu Ala Ala Val Leu His Val Pro Gly Cys Val Pro Cys			
215	220	225	
gag aga gtg ggg aat acg tca cgg tgt tgg gtg cca gtc tcg cca aac			1075
Glu Arg Val Gly Asn Thr Ser Arg Cys Trp Val Pro Val Ser Pro Asn			
230	235	240	245
atg gct gtg cgg cag ccc ggt gcc ctc acg cag ggt ctg cgg acg cac			1123
Met Ala Val Arg Gln Pro Gly Ala Leu Thr Gln Gly Leu Arg Thr His			
250	255	260	
atc gat atg gtt gtg atg tcc gcc acc ttc tgc tct gct ctc tac gtg			1171
Ile Asp Met Val Val Met Ser Ala Thr Phe Cys Ser Ala Leu Tyr Val			
265	270	275	
ggg gac ctc tgt ggc ggg gtg atg ctc gcg gcc cag gtg ttc atc gtc			1219
Gly Asp Leu Cys Gly Gly Val Met Leu Ala Ala Gln Val Phe Ile Val			
280	285	290	
tcg ccg cag tac cac tgg ttt gtg caa gaa tgc aat tgc tcc atc tac			1267
Ser Pro Gln Tyr His Trp Phe Val Gln Glu Cys Asn Cys Ser Ile Tyr			
295	300	305	
cct ggc acc atc act gga cac cgc atg gca tgg gac atg atg atg aac			1315
Pro Gly Thr Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn			
310	315	320	325

tgg tcg ccc acg gcc acc atg atc ctg gcg tac gtg atg cgc gtc ccc 1363

Trp Ser Pro Thr Ala Thr Met Ile Leu Ala Tyr Val Met Arg Val Pro

330

335

340

gag gtc atc ata gac atc gtt agc ggg gct cac tgg ggc gtc atg ttc 1411

Glu Val Ile Ile Asp Ile Val Ser Gly Ala His Trp Gly Val Met Phe

345

350

355

ggc ttg gcc tac ttc tct atg cag gga gcg tgg gcg aag gtc att gtc 1459

Gly Leu Ala Tyr Phe Ser Met Gln Gly Ala Trp Ala Lys Val Ile Val

360

365

370

atc ctt ctg ctg gcc gct ggg gtg gac gcg ggc acc acc acc gtt gga 1507

Ile Leu Leu Leu Ala Ala Gly Val Asp Ala Gly Thr Thr Thr Val Gly

375

380

385

ggc gct gtt gca cgt tcc acc aac gtg att gcc ggc gtg ttc agc cat 1555

Gly Ala Val Ala Arg Ser Thr Asn Val Ile Ala Gly Val Phe Ser His

390

395

400

405

ggc cct cag cag aac att cag ctc att aac acc aac ggc agt tgg cac 1603

Gly Pro Gln Gln Asn Ile Gln Leu Ile Asn Thr Asn Gly Ser Trp His

410

415

420

atc aac cgt act gcc ttg aat tgc aat gac tcc ttg aac acc ggc ttt 1651

Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser Leu Asn Thr Gly Phe

425

430

435

ctc gcg gcc ttg ttc tac acc aac cgc ttt aac tcg tca ggg tgt cca 1699

Leu Ala Ala Leu Phe Tyr Thr Asn Arg Phe Asn Ser Ser Gly Cys Pro

440

445

450

ggg cgc ctg tcc gcc tgc cgc aac atc gag gct ttc cgg ata ggg tgg 1747

Gly Arg Leu Ser Ala Cys Arg Asn Ile Glu Ala Phe Arg Ile Gly Trp

455

460

465

ggc acc cta cag tac gag gat aat gtc acc aat cca gag gat atg agg 1795

Gly Thr Leu Gln Tyr Glu Asp Asn Val Thr Asn Pro Glu Asp Met Arg

470

475

480

485

ccg tac tgc tgg cac tac ccc cca aag ccg tgt ggc gta gtc ccc gcg 1843

Pro Tyr Cys Trp His Tyr Pro Pro Lys Pro Cys Gly Val Val Pro Ala

490

495

500

agg tct gtg tgt ggc cca gtg tac tgt ttc acc ccc agc ccg gta gta 1891

Arg Ser Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val Val

505

510

515

gtg ggc acg acc gac aga cgt gga gtg ccc acc tac aca tgg gga gag 1939

Val Gly Thr Thr Asp Arg Arg Gly Val Pro Thr Tyr Thr Trp Gly Glu

520

525

530

aat gag aca gat gtc ttc cta ctg aac agc acc cga ccg ccg cag ggc 1987

Asn Glu Thr Asp Val Phe Leu Leu Asn Ser Thr Arg Pro Pro Gln Gly

535

540

545

tca tgg ttc ggc tgc acg tgg atg aac tcc act ggt ttc acc aag act 2035

Ser Trp Phe Gly Cys Thr Trp Met Asn Ser Thr Gly Phe Thr Lys Thr  
550 555 560 565

tgt ggc gcg cca cct tgc cgc acc aga gct gac ttc aac gcc agc acg 2083  
Cys Gly Ala Pro Pro Cys Arg Thr Arg Ala Asp Phe Asn Ala Ser Thr  
570 575 580

gac ttg ttg tgc cct acg gat tgt ttt agg aag cat cct gat gcc act 2131  
Asp Leu Leu Cys Pro Thr Asp Cys Phe Arg Lys His Pro Asp Ala Thr  
585 590 595

tat att aag tgt ggt tct ggg ccc tgg ctc aca cca aag tgc ctg gtc 2179  
Tyr Ile Lys Cys Gly Ser Gly Pro Trp Leu Thr Pro Lys Cys Leu Val  
600 605 610

cac tac cct tac aga ctc tgg cat tac ccc tgc aca gtc aat ttt acc 2227  
His Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe Thr  
615 620 625

atc ttc aag ata aga atg tat gta ggg ggg gtt gag cac agg ctc acg 2275  
Ile Phe Lys Ile Arg Met Tyr Val Gly Gly Val Glu His Arg Leu Thr  
630 635 640 645

gcc gca tgc aac ttc act cgt ggg gat cgc tgc gac ttg gag gac agg 2323  
Ala Ala Cys Asn Phe Thr Arg Gly Asp Arg Cys Asp Leu Glu Asp Arg  
650 655 660

gac agg agt cag ctg tct cct ctg ttg cac tct acc acg gaa tgg gcc 2371  
Asp Arg Ser Gln Leu Ser Pro Leu Leu His Ser Thr Thr Glu Trp Ala

665

670

675

atc ctg ccc tgc acc tac tca gac tta ccc gct ttg tca act ggt ctt 2419

Ile Leu Pro Cys Thr Tyr Ser Asp Leu Pro Ala Leu Ser Thr Gly Leu

680

685

690

ctc cac ctt cac cag aac atc gtg gac gta caa tac atg tat ggc ctc 2467

Leu His Leu His Gln Asn Ile Val Asp Val Gln Tyr Met Tyr Gly Leu

695

700

705

tca cct gct atc aca aaa tac gtc gtt cga tgg gag tgg gtg gta ctc 2515

Ser Pro Ala Ile Thr Lys Tyr Val Val Arg Trp Glu Trp Val Val Leu

710

715

720

725

tta ttc ctg ctc tta gcg gac gcc aga gtc tgc gcc tgc ttg tgg atg 2563

Leu Phe Leu Leu Leu Ala Asp Ala Arg Val Cys Ala Cys Leu Trp Met

730

735

740

ctc atc ttg ttg ggc cag gcc gaa gca gca ttg gag aag ttg gtc gtc 2611

Leu Ile Leu Leu Gly Gln Ala Glu Ala Ala Leu Glu Lys Leu Val Val

745

750

755

ttg cac gct gcg agt gcg gct aac tgc cat ggc ctc cta tat ttt gcc 2659

Leu His Ala Ala Ser Ala Ala Asn Cys His Gly Leu Leu Tyr Phe Ala

760

765

770

atc ttc ttc gtg gca gct tgg cac atc agg ggt cgg gtg gtc ccc ttg 2707

Ile Phe Phe Val Ala Ala Trp His Ile Arg Gly Arg Val Val Pro Leu

775

780

785



acc acc tat tgc ctc act ggc cta tgg ccc ttc tgc cta ctg ctc atg 2755  
Thr Thr Tyr Cys Leu Thr Gly Leu Trp Pro Phe Cys Leu Leu Leu Met  
790 795 800 805

gca ctg ccc cgg cag gct tat gcc tat gac gca cct gtg cac gga cag 2803  
Ala Leu Pro Arg Gln Ala Tyr Ala Tyr Asp Ala Pro Val His Gly Gln  
810 815 820

ata ggc gtg ggt ttg ttg ata ttg atc acc ctc ttc aca ctc acc ccg 2851  
Ile Gly Val Gly Leu Leu Ile Leu Ile Thr Leu Phe Thr Leu Thr Pro  
825 830 835

ggg tat aag acc ctc ctc ggc cag tgt ctg tgg tgg ttg tgc tat ctc 2899  
Gly Tyr Lys Thr Leu Leu Gly Gln Cys Leu Trp Trp Leu Cys Tyr Leu  
840 845 850

ctg acc ctg ggg gaa gcc atg att cag gag tgg gta cca ccc atg cag 2947  
Leu Thr Leu Gly Glu Ala Met Ile Gln Glu Trp Val Pro Pro Met Gln  
855 860 865

gtg cgc ggc ggc cgc gat ggc atc gcg tgg gcc gtc act ata ttc tgc 2995  
Val Arg Gly Gly Arg Asp Gly Ile Ala Trp Ala Val Thr Ile Phe Cys  
870 875 880 885

ccg ggt gtg gtg ttt gac att acc aaa tgg ctt ttg gcg ttg ctt ggg 3043  
Pro Gly Val Val Phe Asp Ile Thr Lys Trp Leu Leu Ala Leu Leu Gly  
890 895 900

cct gct tac ctc tta agg gcc gct ttg aca cat gtg ccg tac ttc gtc 3091

Pro Ala Tyr Leu Leu Arg Ala Ala Leu Thr His Val Pro Tyr Phe Val

905

910

915

aga gct cac gct ctg ata agg gta tgc gct ttg gtg aag cag ctc gcg 3139

Arg Ala His Ala Leu Ile Arg Val Cys Ala Leu Val Lys Gln Leu Ala

920

925

930

ggg ggt agg tat gtt cag gtg gcg cta ttg gcc ctt ggc agg tgg act 3187

Gly Gly Arg Tyr Val Gln Val Ala Leu Leu Ala Leu Gly Arg Trp Thr

935

940

945

ggc acc tac atc tat gac cac ctc aca cct atg tcg gac tgg gcc gct 3235

Gly Thr Tyr Ile Tyr Asp His Leu Thr Pro Met Ser Asp Trp Ala Ala

950

955

960

965

agc ggc ctg cgc gac tta gcg gtc gcc gtg gaa ccc atc atc ttc agt 3283

Ser Gly Leu Arg Asp Leu Ala Val Ala Val Glu Pro Ile Ile Phe Ser

970

975

980

ccg atg gag aag aag gtc atc gtc tgg gga gcg gag acg gct gca tgt 3331

Pro Met Glu Lys Lys Val Ile Val Trp Gly Ala Glu Thr Ala Ala Cys

985

990

995

ggg gac att cta cat gga ctt ccc gtg tcc gcc cga ctc ggc cag gag 3379

Gly Asp Ile Leu His Gly Leu Pro Val Ser Ala Arg Leu Gly Gln Glu

1000

1005

1010

atc ctc ctc ggc cca gct gat ggc tac acc tcc aag ggg tgg aag ctc 3427

Ile Leu Leu Gly Pro Ala Asp Gly Tyr Thr Ser Lys Gly Trp Lys Leu

1015

1020

1025

ctt gct ccc atc act gct tat gcc cag caa aca cga ggc ctc ctg ggc 3475

Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu Leu Gly

1030

1035

1040

1045

gcc ata gtg gtg agt atg acg ggg cgt gac agg aca gaa cag gcc ggg 3523

Ala Ile Val Val Ser Met Thr Gly Arg Asp Arg Thr Glu Gln Ala Gly

1050

1055

1060

gaa gtc caa atc ctg tcc aca gtc tct cag tcc ttc ctc gga aca acc 3571

Glu Val Gln Ile Leu Ser Thr Val Ser Gln Ser Phe Leu Gly Thr Thr

1065

1070

1075

atc tcg ggg gtt ttg tgg act gtt tac cac gga gct ggc aac aag act 3619

Ile Ser Gly Val Leu Trp Thr Val Tyr His Gly Ala Gly Asn Lys Thr

1080

1085

1090

cta gcc ggc tta cgg ggt ccg gtc acg cag atg tac tcg agt gct gag 3667

Leu Ala Gly Leu Arg Gly Pro Val Thr Gln Met Tyr Ser Ser Ala Glu

1095

1100

1105

ggg gac ttg gta ggc tgg ccc agc ccc cct ggg acc aag tct ttg gag 3715

Gly Asp Leu Val Gly Trp Pro Ser Pro Pro Gly Thr Lys Ser Leu Glu

1110

1115

1120

1125

ccg tgc aag tgt gga gcc gtc gac cta tat ctg gtc acg cgg aac gct 3763

Pro Cys Lys Cys Gly Ala Val Asp Leu Tyr Leu Val Thr Arg Asn Ala

1130	1135	1140	
gat gtc atc ccg gct cgg aga cgc ggg gac aag cgg gga gca ttg ctc 3811			
Asp Val Ile Pro Ala Arg Arg Arg Gly Asp Lys Arg Gly Ala Leu Leu			
1145	1150	1155	
tcc ccg aga ccc att teg acc ttg aag ggg tcc tcg ggg ggg ccg gtg 3859			
Ser Pro Arg Pro Ile Ser Thr Leu Lys Gly Ser Ser Gly Gly Pro Val			
1160	1165	1170	
ctc tgc cct agg ggc cac gtc gtt ggg ctc ttc cga gca gct gtg tgc 3907			
Leu Cys Pro Arg Gly His Val Val Gly Leu Phe Arg Ala Ala Val Cys			
1175	1180	1185	
tct cgg ggc gtg gcc aaa tcc atc gat ttc atc ccc gtt gag aca ctc 3955			
Ser Arg Gly Val Ala Lys Ser Ile Asp Phe Ile Pro Val Glu Thr Leu			
1190	1195	1200	1205
gac gtt gtt aca agg tct ccc act ttc agt gac aac agc acg cca ccg 4003			
Asp Val Val Thr Arg Ser Pro Thr Phe Ser Asp Asn Ser Thr Pro Pro			
1210	1215	1220	
gct gtg ccc cag acc tat cag gtc ggg tac ttg cat gct cca act ggc 4051			
Ala Val Pro Gln Thr Tyr Gln Val Gly Tyr Leu His Ala Pro Thr Gly			
1225	1230	1235	
agt gga aag agc acc aag gtc cct gtc gcg tat gcc gcc cag ggg tac 4099			
Ser Gly Lys Ser Thr Lys Val Pro Val Ala Tyr Ala Ala Gln Gly Tyr			
1240	1245	1250	

aaa gta cta gtg ctt aac ccc tcg gta gct gcc acc ctg ggg ttt ggg 4147

Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly

1255

1260

1265

gcg tac cta tcc aag gca cat ggc atc aat ccc aac att agg act gga 4195

Ala Tyr Leu Ser Lys Ala His Gly Ile Asn Pro Asn Ile Arg Thr Gly

1270

1275

1280

1285

gtc agg acc gtg atg acc ggg gag gcc atc acg tac tcc aca tat ggc 4243

Val Arg Thr Val Met Thr Gly Glu Ala Ile Thr Tyr Ser Thr Tyr Gly

1290

1295

1300

aaa ttt ctc gcc gat ggg ggc tgc gct agc ggc gcc tat gac atc atc 4291

Lys Phe Leu Ala Asp Gly Gly Cys Ala Ser Gly Ala Tyr Asp Ile Ile

1305

1310

1315

ata tgc gat gaa tgc cac gct gtg gat gct acc tcc att ctc ggc atc 4339

Ile Cys Asp Glu Cys His Ala Val Asp Ala Thr Ser Ile Leu Gly Ile

1320

1325

1330

gga acg gtc ctt gat caa gca gag aca gcc ggg gtc aga cta act gtg 4387

Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Val Arg Leu Thr Val

1335

1340

1345

ctg gct acg gcc aca ccc ccc ggg tca gtg aca acc ccc cat ccc gat 4435

Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Thr Pro His Pro Asp

1350

1355

1360

1365

ata gaa gag gta ggc ctc ggg cgg gag ggt gag atc ccc ttc tat ggg 4483

Ile Glu Glu Val Gly Leu Gly Arg Glu Gly Glu Ile Pro Phe Tyr Gly

1370

1375

1380

agg gcg att ccc cta tcc tgc atc aag gga ggg aga cac ctg att ttc 4531

Arg Ala Ile Pro Leu Ser Cys Ile Lys Gly Gly Arg His Leu Ile Phe

1385

1390

1395

tgc cac tca aag aaa aag tgt gac gag ctc gcg gcg gcc ctt cgg ggc 4579

Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Ala Leu Arg Gly

1400

1405

1410

atg ggc ttg aat gcc gtg gca tac tat aga ggg ttg gac gtc tcc ata 4627

Met Gly Leu Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Ile

1415

1420

1425

ata cca gct cag gga gat gtg gtg gtc gtc gcc acc gac gcc ctc atg 4675

Ile Pro Ala Gln Gly Asp Val Val Val Val Ala Thr Asp Ala Leu Met

1430

1435

1440

1445

acg ggg tac act gga gac ttt gac tcc gtg atc gac tgc aat gta gcg 4723

Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Val Ala

1450

1455

1460

gtc acc caa gct gtc gac ttc agc ctg gac ccc acc ttc act ata acc 4771

Val Thr Gln Ala Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Thr

1465

1470

1475

aca cag act gtc cca caa gac gct gtc tca cgc agt cag cgc cgc ggg 4819

Thr Gln Thr Val Pro Gln Asp Ala Val Ser Arg Ser Gln Arg Arg Gly  
 1480 1485 1490

cgc aca ggt aga gga aga cag ggc act tat agg tat gtt tcc act ggt 4867  
 Arg Thr Gly Arg Gly Arg Gln Gly Thr Tyr Arg Tyr Val Ser Thr Gly  
 1495 1500 1505

gaa cga gcc tca gga atg ttt gac agt gta gtg ctt tgt gag tgc tac 4915  
 Glu Arg Ala Ser Gly Met Phe Asp Ser Val Val Leu Cys Glu Cys Tyr  
 1510 1515 1520 1525

gac gca ggg gct gcg tgg tac gat ctc aca cca gcg gag acc acc gtc 4963  
 Asp Ala Gly Ala Ala Trp Tyr Asp Leu Thr Pro Ala Glu Thr Thr Val  
 1530 1535 1540

agg ctt aga gcg tat ttc aac acg ccc ggc cta ccc gtg tgt caa gac 5011  
 Arg Leu Arg Ala Tyr Phe Asn Thr Pro Gly Leu Pro Val Cys Gln Asp  
 1545 1550 1555

cat ctt gaa ttt tgg gag gca gtt ttc acc ggc ctc aca cac ata gac 5059  
 His Leu Glu Phe Trp Glu Ala Val Phe Thr Gly Leu Thr His Ile Asp  
 1560 1565 1570

gcc cac ttc ctc tcc caa aca aag caa gcg ggg gag aac ttc gcg tac 5107  
 Ala His Phe Leu Ser Gln Thr Lys Gln Ala Gly Glu Asn Phe Ala Tyr  
 1575 1580 1585

cta gta gcc tac caa gct acg gtg tgc gcc aga gcc aag gcc cct ccc 5155  
 Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Lys Ala Pro Pro

1590                      1595                      1600                      1605

ccg tcc tgg gac gcc atg tgg aag tgc ctg gcc cga ctc aag cct acg 5203  
Pro Ser Trp Asp Ala Met Trp Lys Cys Leu Ala Arg Leu Lys Pro Thr

1610                      1615                      1620

ctt gcg ggc ccc aca cct ctc ctg tac cgt ttg ggc cct att acc aat 5251  
Leu Ala Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Pro Ile Thr Asn

1625                      1630                      1635

gag gtc acc ctc aca cac cct ggg acg aag tac atc gcc aca tgc atg 5299  
Glu Val Thr Leu Thr His Pro Gly Thr Lys Tyr Ile Ala Thr Cys Met

1640                      1645                      1650

caa gct gac ctt gag gtc atg acc agc acg tgg gtc cta gct gga gga 5347  
Gln Ala Asp Leu Glu Val Met Thr Ser Thr Trp Val Leu Ala Gly Gly

1655                      1660                      1665

gtc ctg gca gcc gtc gcc gca tat tgc ctg gcg act gga tgc gtt tcc 5395  
Val Leu Ala Ala Val Ala Ala Tyr Cys Leu Ala Thr Gly Cys Val Ser

1670                      1675                      1680                      1685

atc atc ggc cgc ttg cac gtc aac cag cga gtc gtc gtt gcg ccg gat 5443  
Ile Ile Gly Arg Leu His Val Asn Gln Arg Val Val Val Ala Pro Asp

1690                      1695                      1700

aag gag gtc ctg tat gag gct ttt gat gag atg gag gaa tgc gcc tct 5491  
Lys Glu Val Leu Tyr Glu Ala Phe Asp Glu Met Glu Glu Cys Ala Ser

1705                      1710                      1715



agg gcg gct ctc atc gaa gag ggg cag cgg ata gcc gag atg ttg aag 5539

Arg Ala Ala Leu Ile Glu Glu Gly Gln Arg Ile Ala Glu Met Leu Lys

1720

1725

1730

tcc aag atc caa ggc ttg ctg cag cag gcc tct aag cag gcc cag gac 5587

Ser Lys Ile Gln Gly Leu Leu Gln Gln Ala Ser Lys Gln Ala Gln Asp

1735

1740

1745

ata caa ccc gct atg cag gct tca tgg ccc aaa gtg gaa caa ttt tgg 5635

Ile Gln Pro Ala Met Gln Ala Ser Trp Pro Lys Val Glu Gln Phe Trp

1750

1755

1760

1765

gcc aga cac atg tgg aac ttc att agc ggc atc caa tac ctc gca gga 5683

Ala Arg His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala Gly

1770

1775

1780

ttg tca aca ctg cca ggg aac ccc gcg gtg gct tcc atg atg gca ttc 5731

Leu Ser Thr Leu Pro Gly Asn Pro Ala Val Ala Ser Met Met Ala Phe

1785

1790

1795

agt gcc gcc ctc acc agt ccg ttg tcg acc agt acc acc atc ctt ctc 5779

Ser Ala Ala Leu Thr Ser Pro Leu Ser Thr Ser Thr Thr Ile Leu Leu

1800

1805

1810

aac atc atg gga ggc tgg tta gcg tcc cag atc gca cca ccc gcg ggg 5827

Asn Ile Met Gly Gly Trp Leu Ala Ser Gln Ile Ala Pro Pro Ala Gly

1815

1820

1825

gcc acc ggc ttt gtc gtc agt ggc ctg gtg ggg gct gcc gtg ggc agc 5875  
Ala Thr Gly Phe Val Val Ser Gly Leu Val Gly Ala Ala Val Gly Ser  
1830 1835 1840 1845

ata ggc ctg ggt aag gtg ctg gtg gac atc ctg gca gga tat ggt gcg 5923  
Ile Gly Leu Gly Lys Val Leu Val Asp Ile Leu Ala Gly Tyr Gly Ala  
1850 1855 1860

ggc att tcg ggg gcc ctc gtc gca ttc aag atc atg tct ggc gag aag 5971  
Gly Ile Ser Gly Ala Leu Val Ala Phe Lys Ile Met Ser Gly Glu Lys  
1865 1870 1875

ccc tct atg gaa gat gtc atc aat cta ctg cct ggg atc ctg tct ccg 6019  
Pro Ser Met Glu Asp Val Ile Asn Leu Leu Pro Gly Ile Leu Ser Pro  
1880 1885 1890

gga gcc ctg gtg gtg ggg gtc atc tgc gcg gcc att ctg cgc cgc cac 6067  
Gly Ala Leu Val Val Gly Val Ile Cys Ala Ala Ile Leu Arg Arg His  
1895 1900 1905

gtg gga ccg ggg gag ggc gcg gtc caa tgg atg aac agg ctt att gcc 6115  
Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile Ala  
1910 1915 1920 1925

ttt gct tcc aga gga aac cac gtc gcc cct act cac tac gtg acg gag 6163  
Phe Ala Ser Arg Gly Asn His Val Ala Pro Thr His Tyr Val Thr Glu  
1930 1935 1940

tcg gat gcg tcg cag cgt gtg acc caa cta ctt ggc tct ctt act ata 6211

Ser Asp Ala Ser Gln Arg Val Thr Gln Leu Leu Gly Ser Leu Thr Ile  
 1945 1950 1955

acc agc cta ctc aga aga ctc cac aat tgg ata act gag gac tgc ccc 6259  
 Thr Ser Leu Leu Arg Arg Leu His Asn Trp Ile Thr Glu Asp Cys Pro  
 1960 1965 1970

atc cca tgc tcc gga tcc tgg ctc cgc gac gtg tgg gac tgg gtt tgc 6307  
 Ile Pro Cys Ser Gly Ser Trp Leu Arg Asp Val Trp Asp Trp Val Cys  
 1975 1980 1985

acc atc ttg aca gac ttc aaa aat tgg ctg acc tct aaa ttg ttc ccc 6355  
 Thr Ile Leu Thr Asp Phe Lys Asn Trp Leu Thr Ser Lys Leu Phe Pro  
 1990 1995 2000 2005

aag ctg ccc ggc ctc ccc ttc atc tct tgt caa aag ggg tac aag ggt 6403  
 Lys Leu Pro Gly Leu Pro Phe Ile Ser Cys Gln Lys Gly Tyr Lys Gly  
 2010 2015 2020

gtg tgg gcc ggc act ggc atc atg acc acg cgc tgc cct tgc ggc gcc 6451  
 Val Trp Ala Gly Thr Gly Ile Met Thr Thr Arg Cys Pro Cys Gly Ala  
 2025 2030 2035

aac atc tct ggc aat gtc cgc ctg ggc tct atg agg atc aca ggg cct 6499  
 Asn Ile Ser Gly Asn Val Arg Leu Gly Ser Met Arg Ile Thr Gly Pro  
 2040 2045 2050

aaa acc tgc atg aac acc tgg cag ggg acc ttt cct atc aat tgc tac 6547  
 Lys Thr Cys Met Asn Thr Trp Gln Gly Thr Phe Pro Ile Asn Cys Tyr

2055	2060	2065	
acg gag ggc cag tgc gcg ccg aaa ccc ccc acg aac tac aag acc gcc			6595
Thr Glu Gly Gln Cys Ala Pro Lys Pro Pro Thr Asn Tyr Lys Thr Ala			
2070	2075	2080	2085
atc tgg agg gtg gcg gcc tcg gag tac gcg gag gtg acg cag cat ggg			6643
Ile Trp Arg Val Ala Ala Ser Glu Tyr Ala Glu Val Thr Gln His Gly			
2090	2095	2100	
tcg tac tcc tat gta aca gga ctg acc act gac aat ctg aaa att cct			6691
Ser Tyr Ser Tyr Val Thr Gly Leu Thr Thr Asp Asn Leu Lys Ile Pro			
2105	2110	2115	
tgc caa cta cct tct cca gag ttt ttc tcc tgg gtg gac ggt gtg cag			6739
Cys Gln Leu Pro Ser Pro Glu Phe Phe Ser Trp Val Asp Gly Val Gln			
2120	2125	2130	
atc cat agg ttt gca ccc aca cca aag ccg ttt ttc cgg gat gag gtc			6787
Ile His Arg Phe Ala Pro Thr Pro Lys Pro Phe Phe Arg Asp Glu Val			
2135	2140	2145	
tcg ttc tgc gtt ggg ctt aat tcc tat gct gtc ggg tcc cag ctt ccc			6835
Ser Phe Cys Val Gly Leu Asn Ser Tyr Ala Val Gly Ser Gln Leu Pro			
2150	2155	2160	2165
tgt gaa cct gag ccc gac gca gac gta ttg agg tcc atg cta aca gat			6883
Cys Glu Pro Glu Pro Asp Ala Asp Val Leu Arg Ser Met Leu Thr Asp			
2170	2175	2180	

ccg ccc cac atc acg gcg gag act gcg gcg cgg cgc ttg gca cgg gga 6931

Pro Pro His Ile Thr Ala Glu Thr Ala Ala Arg Arg Leu Ala Arg Gly

2185

2190

2195

tca cct cca tct gag gcg agc tcc tca gtg agc cag cta tca gca ccg 6979

Ser Pro Pro Ser Glu Ala Ser Ser Ser Val Ser Gln Leu Ser Ala Pro

2200

2205

2210

tcg ctg cgg gcc acc tgc acc acc cac agc aac acc tat gac gtg gac 7027

Ser Leu Arg Ala Thr Cys Thr Thr His Ser Asn Thr Tyr Asp Val Asp

2215

2220

2225

atg gtc gat gcc aac ctg ctc atg gag ggc ggt gtg gct cag aca gag 7075

Met Val Asp Ala Asn Leu Leu Met Glu Gly Gly Val Ala Gln Thr Glu

2230

2235

2240

2245

cct gag tcc agg gtg ccc gtt ctg gac ttt ctc gag cca atg gcc gag 7123

Pro Glu Ser Arg Val Pro Val Leu Asp Phe Leu Glu Pro Met Ala Glu

2250

2255

2260

gaa gag agc gac ctt gag ccc tca ata cca tcg gag tgc atg ctc ccc 7171

Glu Glu Ser Asp Leu Glu Pro Ser Ile Pro Ser Glu Cys Met Leu Pro

2265

2270

2275

agg agc ggg ttt cca cgg gcc tta ccg gct tgg gca cgg cct gac tac 7219

Arg Ser Gly Phe Pro Arg Ala Leu Pro Ala Trp Ala Arg Pro Asp Tyr

2280

2285

2290

aac ccg ccg ctc gtg gaa tcg tgg agg agg cca gat tac caa ccg ccc 7267

Asn Pro Pro Leu Val Glu Ser Trp Arg Arg Pro Asp Tyr Gln Pro Pro

2295

2300

2305

acc gtt gct ggt tgt gct ctc ccc ccc ccc aag aag gcc ccg acg cct 7315

Thr Val Ala Gly Cys Ala Leu Pro Pro Pro Lys Lys Ala Pro Thr Pro

2310

2315

2320

2325

ccc cca agg aga cgc cgg aca gtg ggt ctg agc gag agc acc ata tca 7363

Pro Pro Arg Arg Arg Arg Thr Val Gly Leu Ser Glu Ser Thr Ile Ser

2330

2335

2340

gaa gcc ctc cag caa ctg gcc atc aag acc ttt ggc cag ccc ccc tcg 7411

Glu Ala Leu Gln Gln Leu Ala Ile Lys Thr Phe Gly Gln Pro Pro Ser

2345

2350

2355

agc ggt gat gca ggc tcg tcc acg ggg gcg ggc gcc gcc gaa tcc ggc 7459

Ser Gly Asp Ala Gly Ser Ser Thr Gly Ala Gly Ala Ala Glu Ser Gly

2360

2365

2370

ggt ccg acg tcc cct ggt gag ccg gcc ccc tca gag aca ggt tcc gcc 7507

Gly Pro Thr Ser Pro Gly Glu Pro Ala Pro Ser Glu Thr Gly Ser Ala

2375

2380

2385

tcc tct atg ccc ccc ctc gag ggg gag cct gga gat ccg gac ctg gag 7555

Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu Glu

2390

2395

2400

2405

tct gat cag gta gag ctt caa cct ccc ccc cag ggg ggg ggg gta gct 7603

Ser Asp Gln Val Glu Leu Gln Pro Pro Pro Gln Gly Gly Gly Val Ala

2410

2415

2420

ccc ggt tgc ggc tgc ggg tct tgg tct act tgc tcc gag gag gac gat 7651

Pro Gly Ser Gly Ser Gly Ser Trp Ser Thr Cys Ser Glu Glu Asp Asp

2425

2430

2435

acc acc gtg tgc tgc tcc atg tca tac tcc tgg acc ggg gct cta ata 7699

Thr Thr Val Cys Cys Ser Met Ser Tyr Ser Trp Thr Gly Ala Leu Ile

2440

2445

2450

act ccc tgt agc ccc gaa gag gaa aag ttg cca atc aac cct ttg agt 7747

Thr Pro Cys Ser Pro Glu Glu Glu Lys Leu Pro Ile Asn Pro Leu Ser

2455

2460

2465

aac tgc ctg ttg cga tac cat aac aag gtg tac tgt aca aca tca aag 7795

Asn Ser Leu Leu Arg Tyr His Asn Lys Val Tyr Cys Thr Thr Ser Lys

2470

2475

2480

2485

agc gcc tca cag agg gct aaa aag gta act ttt gac agg acg caa gtg 7843

Ser Ala Ser Gln Arg Ala Lys Lys Val Thr Phe Asp Arg Thr Gln Val

2490

2495

2500

ctc gac gcc cat tat gac tca gtc tta aag gac atc aag cta gcg gct 7891

Leu Asp Ala His Tyr Asp Ser Val Leu Lys Asp Ile Lys Leu Ala Ala

2505

2510

2515

tcc aag gtc agc gca agg ctc ctc acc ttg gag gag gcg tgc cag ttg 7939

Ser Lys Val Ser Ala Arg Leu Leu Thr Leu Glu Glu Ala Cys Gln Leu

2520	2525	2530	
act cca ccc cat tct gca aga tcc aag tat gga ttc ggg gcc aag gag 7987			
Thr Pro Pro His Ser Ala Arg Ser Lys Tyr Gly Phe Gly Ala Lys Glu			
2535	2540	2545	
gtc cgc agc ttg tcc ggg agg gcc gtt aac cac atc aag tcc gtg tgg 8035			
Val Arg Ser Leu Ser Gly Arg Ala Val Asn His Ile Lys Ser Val Trp			
2550	2555	2560	2565
aag gac ctc ctg gaa gac cca caa aca cca att ccc aca acc atc atg 8083			
Lys Asp Leu Leu Glu Asp Pro Gln Thr Pro Ile Pro Thr Thr Ile Met			
2570	2575	2580	
gcc aaa aat gag gtg ttc tgc gtg gac ccc gcc aag ggg ggt aag aaa 8131			
Ala Lys Asn Glu Val Phe Cys Val Asp Pro Ala Lys Gly Gly Lys Lys			
2585	2590	2595	
cca gct cgc ctc atc gtt tac cct gac ctc ggc gtc cgg gtc tgc gag 8179			
Pro Ala Arg Leu Ile Val Tyr Pro Asp Leu Gly Val Arg Val Cys Glu			
2600	2605	2610	
aaa atg gcc ctc tat gac att aca caa aag ctt cct cag gcg gta atg 8227			
Lys Met Ala Leu Tyr Asp Ile Thr Gln Lys Leu Pro Gln Ala Val Met			
2615	2620	2625	
gga gct tcc tat ggc ttc cag tac tcc cct gcc caa cgg gtg gag tat 8275			
Gly Ala Ser Tyr Gly Phe Gln Tyr Ser Pro Ala Gln Arg Val Glu Tyr			
2630	2635	2640	2645



ctc ttg aaa gca tgg gcg gaa aag aag gac ccc atg ggt ttt tcg tat 8323  
Leu Leu Lys Ala Trp Ala Glu Lys Lys Asp Pro Met Gly Phe Ser Tyr

2650

2655

2660

gat acc cga tgc ttc gac tca acc gtc act gag aga gac atc agg acc 8371  
Asp Thr Arg Cys Phe Asp Ser Thr Val Thr Glu Arg Asp Ile Arg Thr

2665

2670

2675

gag gag tcc ata tac cag gcc tgc tcc ctg ccc gag gag gcc cgc act 8419  
Glu Glu Ser Ile Tyr Gln Ala Cys Ser Leu Pro Glu Glu Ala Arg Thr

2680

2685

2690

gcc ata cac tcg ctg act gag aga ctt tac gta gga ggg ccc atg ttc 8467  
Ala Ile His Ser Leu Thr Glu Arg Leu Tyr Val Gly Gly Pro Met Phe

2695

2700

2705

aac agc aag ggt caa acc tgc ggt tac aga cgt tgc cgc gcc agc ggg 8515  
Asn Ser Lys Gly Gln Thr Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly

2710

2715

2720

2725

gtg cta acc act agc atg ggt aac acc atc aca tgc tat gtg aaa gcc 8563  
Val Leu Thr Thr Ser Met Gly Asn Thr Ile Thr Cys Tyr Val Lys Ala

2730

2735

2740

cta gcg gcc tgc aag gct gcg ggg ata gtt gcg ccc aca atg ctg gta 8611  
Leu Ala Ala Cys Lys Ala Ala Gly Ile Val Ala Pro Thr Met Leu Val

2745

2750

2755

tgc ggc gat gac cta gta gtc atc tca gaa agc cag ggg act gag gag 8659  
Cys Gly Asp Asp Leu Val Val Ile Ser Glu Ser Gln Gly Thr Glu Glu

2760

2765

2770

gac gag cgg aac ctg aga gcc ttc acg gag gcc atg acc agg tac tct 8707  
Asp Glu Arg Asn Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser

2775

2780

2785

gcc cct cct ggt gat ccc ccc aga ccg gaa tat gac ctg gag cta ata 8755  
Ala Pro Pro Gly Asp Pro Pro Arg Pro Glu Tyr Asp Leu Glu Leu Ile  
2790 2795 2800 2805

aca tcc tgt tcc tca aat gtg tct gtg gcg ttg ggc ccg cgg ggc cgc 8803  
Thr Ser Cys Ser Ser Asn Val Ser Val Ala Leu Gly Pro Arg Gly Arg

2810

2815

2820

cgc aga tac tac ctg acc aga gac cca acc act cca ctc gcc cgg gct 8851  
Arg Arg Tyr Tyr Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala

2825

2830

2835

gcc tgg gaa aca gtt aga cac tcc cct atc aat tca tgg ctg gga aac 8899  
Ala Trp Glu Thr Val Arg His Ser Pro Ile Asn Ser Trp Leu Gly Asn

2840

2845

2850

atc atc cag tat gct cca acc ata tgg gtt cgc atg gtc cta atg aca 8947  
Ile Ile Gln Tyr Ala Pro Thr Ile Trp Val Arg Met Val Leu Met Thr

2855

2860

2865

cac ttc ttc tcc att ctc atg gtc caa gac acc ctg gac cag aac ctc 8995

His Phe Phe Ser Ile Leu Met Val Gln Asp Thr Leu Asp Gln Asn Leu  
2870 2875 2880 2885

aac ttt gag atg tat gga tca gta tac tcc gtg aat cct ttg gac ctt 9043  
Asn Phe Glu Met Tyr Gly Ser Val Tyr Ser Val Asn Pro Leu Asp Leu  
2890 2895 2900

cca gcc ata att gag agg tta cac ggg ctt gac gcc ttt tct atg cac 9091  
Pro Ala Ile Ile Glu Arg Leu His Gly Leu Asp Ala Phe Ser Met His  
2905 2910 2915

aca tac tct cac cac gaa ctg acg cgg gtg gct tca gcc ctc aga aaa 9139  
Thr Tyr Ser His His Glu Leu Thr Arg Val Ala Ser Ala Leu Arg Lys  
2920 2925 2930

ctt ggg gcg cca ccc ctc agg gtg tgg aag agt cgg gct cgc gca gtc 9187  
Leu Gly Ala Pro Pro Leu Arg Val Trp Lys Ser Arg Ala Arg Ala Val  
2935 2940 2945

agg gcg tcc ctc atc tcc cgt gga ggg aaa gcg gcc gtt tgc ggc cga 9235  
Arg Ala Ser Leu Ile Ser Arg Gly Gly Lys Ala Ala Val Cys Gly Arg  
2950 2955 2960 2965

tat ctc ttc aat tgg gcg gtg aag acc aag ctc aaa ctc act cca ttg 9283  
Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu Lys Leu Thr Pro Leu  
2970 2975 2980

ccg gag gcg cgc cta ctg gac tta tcc agt tgg ttc acc gtc ggc gcc 9331  
Pro Glu Ala Arg Leu Leu Asp Leu Ser Ser Trp Phe Thr Val Gly Ala

2985

2990

2995

ggc ggg ggc gac att ttt cac agc gtg tcg cgc gcc cga ccc cgc tca 9379

Gly Gly Gly Asp Ile Phe His Ser Val Ser Arg Ala Arg Pro Arg Ser

3000

3005

3010

tta ctc ttc ggc cta ctc cta ctt ttc gta ggg gta ggc ctc ttc cta 9427

Leu Leu Phe Gly Leu Leu Leu Leu Phe Val Gly Val Gly Leu Phe Leu

3015

3020

3025

ctc ccc gct cgg tag agcggcacac actaggtaca ctccatagct aactgttcct 9482

Leu Pro Ala Arg

3030

tttttttttt tttttttttt tttttttttt tttttttttt ttcttttttt tttttttccc 9542

tctttcttcc cttctcatct tattctactt tctttcttgg tggctccatc ttagccctag 9602

tcacggctag ctgtgaaagg tccgtgagcc gcatgactgc agagagtgcc gtaactggtc 9662

tctctgcaga tcatgt

9678

&lt;210&gt; 4

&lt;211&gt; 3033

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 4

Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn  
 1 5 10 15  
 Arg Arg Pro Glu Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly  
 20 25 30  
 Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Thr  
 35 40 45  
 Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro  
 50 55 60  
 Ile Pro Lys Asp Arg Arg Ser Thr Gly Lys Ala Trp Gly Lys Pro Gly  
 65 70 75 80  
 Arg Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp  
 85 90 95  
 Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro  
 100 105 110  
 Arg His Arg Ser Arg Asn Val Gly Lys Val Ile Asp Thr Leu Thr Cys  
 115 120 125  
 Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Val Val Gly Ala Pro Leu  
 130 135 140  
 Ser Gly Ala Ala Arg Ala Val Ala His Gly Val Arg Val Leu Glu Asp  
 145 150 155 160  
 Gly Val Asn Tyr Ala Thr Gly Asn Leu Pro Gly Phe Pro Phe Ser Ile  
 165 170 175  
 Phe Leu Leu Ala Leu Leu Ser Cys Ile Thr Val Pro Val Ser Ala Ala  
 180 185 190  
 Gln Val Lys Asn Thr Ser Ser Ser Tyr Met Val Thr Asn Asp Cys Ser  
 195 200 205  
 Asn Asp Ser Ile Thr Trp Gln Leu Glu Ala Ala Val Leu His Val Pro  
 210 215 220  
 Gly Cys Val Pro Cys Glu Arg Val Gly Asn Thr Ser Arg Cys Trp Val

225                      230                      235                      240  
Pro Val Ser Pro Asn Met Ala Val Arg Gln Pro Gly Ala Leu Thr Gln  
245                      250                      255  
Gly Leu Arg Thr His Ile Asp Met Val Val Met Ser Ala Thr Phe Cys  
260                      265                      270  
Ser Ala Leu Tyr Val Gly Asp Leu Cys Gly Gly Val Met Leu Ala Ala  
275                      280                      285  
Gln Val Phe Ile Val Ser Pro Gln Tyr His Trp Phe Val Gln Glu Cys  
290                      295                      300  
Asn Cys Ser Ile Tyr Pro Gly Thr Ile Thr Gly His Arg Met Ala Trp  
305                      310                      315                      320  
Asp Met Met Met Asn Trp Ser Pro Thr Ala Thr Met Ile Leu Ala Tyr  
325                      330                      335  
Val Met Arg Val Pro Glu Val Ile Ile Asp Ile Val Ser Gly Ala His  
340                      345                      350  
Trp Gly Val Met Phe Gly Leu Ala Tyr Phe Ser Met Gln Gly Ala Trp  
355                      360                      365  
Ala Lys Val Ile Val Ile Leu Leu Leu Ala Ala Gly Val Asp Ala Gly  
370                      375                      380  
Thr Thr Thr Val Gly Gly Ala Val Ala Arg Ser Thr Asn Val Ile Ala  
385                      390                      395                      400  
Gly Val Phe Ser His Gly Pro Gln Gln Asn Ile Gln Leu Ile Asn Thr  
405                      410                      415  
Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser  
420                      425                      430  
Leu Asn Thr Gly Phe Leu Ala Ala Leu Phe Tyr Thr Asn Arg Phe Asn  
435                      440                      445  
Ser Ser Gly Cys Pro Gly Arg Leu Ser Ala Cys Arg Asn Ile Glu Ala  
450                      455                      460

Phe Arg Ile Gly Trp Gly Thr Leu Gln Tyr Glu Asp Asn Val Thr Asn  
465 470 475 480  
Pro Glu Asp Met Arg Pro Tyr Cys Trp His Tyr Pro Pro Lys Pro Cys  
485 490 495  
Gly Val Val Pro Ala Arg Ser Val Cys Gly Pro Val Tyr Cys Phe Thr  
500 505 510  
Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Arg Gly Val Pro Thr  
515 520 525  
Tyr Thr Trp Gly Glu Asn Glu Thr Asp Val Phe Leu Leu Asn Ser Thr  
530 535 540  
Arg Pro Pro Gln Gly Ser Trp Phe Gly Cys Thr Trp Met Asn Ser Thr  
545 550 555 560  
Gly Phe Thr Lys Thr Cys Gly Ala Pro Pro Cys Arg Thr Arg Ala Asp  
565 570 575  
Phe Asn Ala Ser Thr Asp Leu Leu Cys Pro Thr Asp Cys Phe Arg Lys  
580 585 590  
His Pro Asp Ala Thr Tyr Ile Lys Cys Gly Ser Gly Pro Trp Leu Thr  
595 600 605  
Pro Lys Cys Leu Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys  
610 615 620  
Thr Val Asn Phe Thr Ile Phe Lys Ile Arg Met Tyr Val Gly Gly Val  
625 630 635 640  
Glu His Arg Leu Thr Ala Ala Cys Asn Phe Thr Arg Gly Asp Arg Cys  
645 650 655  
Asp Leu Glu Asp Arg Asp Arg Ser Gln Leu Ser Pro Leu Leu His Ser  
660 665 670  
Thr Thr Glu Trp Ala Ile Leu Pro Cys Thr Tyr Ser Asp Leu Pro Ala  
675 680 685  
Leu Ser Thr Gly Leu Leu His Leu His Gln Asn Ile Val Asp Val Gln

690                      695                      700  
 Tyr Met Tyr Gly Leu Ser Pro Ala Ile Thr Lys Tyr Val Val Arg Trp  
 705                      710                      715                      720  
 Glu Trp Val Val Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Val Cys  
                     725                      730                      735  
 Ala Cys Leu Trp Met Leu Ile Leu Leu Gly Gln Ala Glu Ala Ala Leu  
                     740                      745                      750  
 Glu Lys Leu Val Val Leu His Ala Ala Ser Ala Ala Asn Cys His Gly  
                     755                      760                      765  
 Leu Leu Tyr Phe Ala Ile Phe Phe Val Ala Ala Trp His Ile Arg Gly  
                     770                      775                      780  
 Arg Val Val Pro Leu Thr Thr Tyr Cys Leu Thr Gly Leu Trp Pro Phe  
 785                      790                      795                      800  
 Cys Leu Leu Leu Met Ala Leu Pro Arg Gln Ala Tyr Ala Tyr Asp Ala  
                     805                      810                      815  
 Pro Val His Gly Gln Ile Gly Val Gly Leu Leu Ile Leu Ile Thr Leu  
                     820                      825                      830  
 Phe Thr Leu Thr Pro Gly Tyr Lys Thr Leu Leu Gly Gln Cys Leu Trp  
                     835                      840                      845  
 Trp Leu Cys Tyr Leu Leu Thr Leu Gly Glu Ala Met Ile Gln Glu Trp  
                     850                      855                      860  
 Val Pro Pro Met Gln Val Arg Gly Gly Arg Asp Gly Ile Ala Trp Ala  
 865                      870                      875                      880  
 Val Thr Ile Phe Cys Pro Gly Val Val Phe Asp Ile Thr Lys Trp Leu  
                     885                      890                      895  
 Leu Ala Leu Leu Gly Pro Ala Tyr Leu Leu Arg Ala Ala Leu Thr His  
                     900                      905                      910  
 Val Pro Tyr Phe Val Arg Ala His Ala Leu Ile Arg Val Cys Ala Leu  
                     915                      920                      925



Val Lys Gln Leu Ala Gly Gly Arg Tyr Val Gln Val Ala Leu Leu Ala  
930 935 940

Leu Gly Arg Trp Thr Gly Thr Tyr Ile Tyr Asp His Leu Thr Pro Met  
945 950 955 960

Ser Asp Trp Ala Ala Ser Gly Leu Arg Asp Leu Ala Val Ala Val Glu  
965 970 975

Pro Ile Ile Phe Ser Pro Met Glu Lys Lys Val Ile Val Trp Gly Ala  
980 985 990

Glu Thr Ala Ala Cys Gly Asp Ile Leu His Gly Leu Pro Val Ser Ala  
995 1000 1005

Arg Leu Gly Gln Glu Ile Leu Leu Gly Pro Ala Asp Gly Tyr Thr Ser  
1010 1015 1020

Lys Gly Trp Lys Leu Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr  
1025 1030 1035 1040

Arg Gly Leu Leu Gly Ala Ile Val Val Ser Met Thr Gly Arg Asp Arg  
1045 1050 1055

Thr Glu Gln Ala Gly Glu Val Gln Ile Leu Ser Thr Val Ser Gln Ser  
1060 1065 1070

Phe Leu Gly Thr Thr Ile Ser Gly Val Leu Trp Thr Val Tyr His Gly  
1075 1080 1085

Ala Gly Asn Lys Thr Leu Ala Gly Leu Arg Gly Pro Val Thr Gln Met  
1090 1095 1100

Tyr Ser Ser Ala Glu Gly Asp Leu Val Gly Trp Pro Ser Pro Pro Gly  
1105 1110 1115 1120

Thr Lys Ser Leu Glu Pro Cys Lys Cys Gly Ala Val Asp Leu Tyr Leu  
1125 1130 1135

Val Thr Arg Asn Ala Asp Val Ile Pro Ala Arg Arg Arg Gly Asp Lys  
1140 1145 1150

Arg Gly Ala Leu Leu Ser Pro Arg Pro Ile Ser Thr Leu Lys Gly Ser

1155	1160	1165
Ser Gly Gly Pro Val Leu Cys Pro Arg Gly His Val Val Gly Leu Phe		
1170	1175	1180
Arg Ala Ala Val Cys Ser Arg Gly Val Ala Lys Ser Ile Asp Phe Ile		
1185	1190	1195
Pro Val Glu Thr Leu Asp Val Val Thr Arg Ser Pro Thr Phe Ser Asp		1200
1205	1210	1215
Asn Ser Thr Pro Pro Ala Val Pro Gln Thr Tyr Gln Val Gly Tyr Leu		
1220	1225	1230
His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro Val Ala Tyr		
1235	1240	1245
Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala		
1250	1255	1260
Thr Leu Gly Phe Gly Ala Tyr Leu Ser Lys Ala His Gly Ile Asn Pro		
1265	1270	1275
Asn Ile Arg Thr Gly Val Arg Thr Val Met Thr Gly Glu Ala Ile Thr		1280
1285	1290	1295
Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys Ala Ser Gly		
1300	1305	1310
Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ala Val Asp Ala Thr		
1315	1320	1325
Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly		
1330	1335	1340
Val Arg Leu Thr Val Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr		
1345	1350	1355
Thr Pro His Pro Asp Ile Glu Glu Val Gly Leu Gly Arg Glu Gly Glu		1360
1365	1370	1375
Ile Pro Phe Tyr Gly Arg Ala Ile Pro Leu Ser Cys Ile Lys Gly Gly		
1380	1385	1390

Arg His Leu Ile Phe Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala  
1395 1400 1405  
Ala Ala Leu Arg Gly Met Gly Leu Asn Ala Val Ala Tyr Tyr Arg Gly  
1410 1415 1420  
Leu Asp Val Ser Ile Ile Pro Ala Gln Gly Asp Val Val Val Val Ala  
1425 1430 1435 1440  
Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile  
1445 1450 1455  
Asp Cys Asn Val Ala Val Thr Gln Ala Val Asp Phe Ser Leu Asp Pro  
1460 1465 1470  
Thr Phe Thr Ile Thr Thr Gln Thr Val Pro Gln Asp Ala Val Ser Arg  
1475 1480 1485  
Ser Gln Arg Arg Gly Arg Thr Gly Arg Gly Arg Gln Gly Thr Tyr Arg  
1490 1495 1500  
Tyr Val Ser Thr Gly Glu Arg Ala Ser Gly Met Phe Asp Ser Val Val  
1505 1510 1515 1520  
Leu Cys Glu Cys Tyr Asp Ala Gly Ala Ala Trp Tyr Asp Leu Thr Pro  
1525 1530 1535  
Ala Glu Thr Thr Val Arg Leu Arg Ala Tyr Phe Asn Thr Pro Gly Leu  
1540 1545 1550  
Pro Val Cys Gln Asp His Leu Glu Phe Trp Glu Ala Val Phe Thr Gly  
1555 1560 1565  
Leu Thr His Ile Asp Ala His Phe Leu Ser Gln Thr Lys Gln Ala Gly  
1570 1575 1580  
Glu Asn Phe Ala Tyr Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg  
1585 1590 1595 1600  
Ala Lys Ala Pro Pro Pro Ser Trp Asp Ala Met Trp Lys Cys Leu Ala  
1605 1610 1615  
Arg Leu Lys Pro Thr Leu Ala Gly Pro Thr Pro Leu Leu Tyr Arg Leu

1620 1625 1630  
Gly Pro Ile Thr Asn Glu Val Thr Leu Thr His Pro Gly Thr Lys Tyr  
1635 1640 1645  
Ile Ala Thr Cys Met Gln Ala Asp Leu Glu Val Met Thr Ser Thr Trp  
1650 1655 1660  
Val Leu Ala Gly Gly Val Leu Ala Ala Val Ala Ala Tyr Cys Leu Ala  
1665 1670 1675 1680  
Thr Gly Cys Val Ser Ile Ile Gly Arg Leu His Val Asn Gln Arg Val  
1685 1690 1695  
Val Val Ala Pro Asp Lys Glu Val Leu Tyr Glu Ala Phe Asp Glu Met  
1700 1705 1710  
Glu Glu Cys Ala Ser Arg Ala Ala Leu Ile Glu Glu Gly Gln Arg Ile  
1715 1720 1725  
Ala Glu Met Leu Lys Ser Lys Ile Gln Gly Leu Leu Gln Gln Ala Ser  
1730 1735 1740  
Lys Gln Ala Gln Asp Ile Gln Pro Ala Met Gln Ala Ser Trp Pro Lys  
1745 1750 1755 1760  
Val Glu Gln Phe Trp Ala Arg His Met Trp Asn Phe Ile Ser Gly Ile  
1765 1770 1775  
Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn Pro Ala Val Ala  
1780 1785 1790  
Ser Met Met Ala Phe Ser Ala Ala Leu Thr Ser Pro Leu Ser Thr Ser  
1795 1800 1805  
Thr Thr Ile Leu Leu Asn Ile Met Gly Gly Trp Leu Ala Ser Gln Ile  
1810 1815 1820  
Ala Pro Pro Ala Gly Ala Thr Gly Phe Val Val Ser Gly Leu Val Gly  
1825 1830 1835 1840  
Ala Ala Val Gly Ser Ile Gly Leu Gly Lys Val Leu Val Asp Ile Leu  
1845 1850 1855

Ala Gly Tyr Gly Ala Gly Ile Ser Gly Ala Leu Val Ala Phe Lys Ile  
1860 1865 1870  
Met Ser Gly Glu Lys Pro Ser Met Glu Asp Val Ile Asn Leu Leu Pro  
1875 1880 1885  
Gly Ile Leu Ser Pro Gly Ala Leu Val Val Gly Val Ile Cys Ala Ala  
1890 1895 1900  
Ile Leu Arg Arg His Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met  
1905 1910 1915 1920  
Asn Arg Leu Ile Ala Phe Ala Ser Arg Gly Asn His Val Ala Pro Thr  
1925 1930 1935  
His Tyr Val Thr Glu Ser Asp Ala Ser Gln Arg Val Thr Gln Leu Leu  
1940 1945 1950  
Gly Ser Leu Thr Ile Thr Ser Leu Leu Arg Arg Leu His Asn Trp Ile  
1955 1960 1965  
Thr Glu Asp Cys Pro Ile Pro Cys Ser Gly Ser Trp Leu Arg Asp Val  
1970 1975 1980  
Trp Asp Trp Val Cys Thr Ile Leu Thr Asp Phe Lys Asn Trp Leu Thr  
1985 1990 1995 2000  
Ser Lys Leu Phe Pro Lys Leu Pro Gly Leu Pro Phe Ile Ser Cys Gln  
2005 2010 2015  
Lys Gly Tyr Lys Gly Val Trp Ala Gly Thr Gly Ile Met Thr Thr Arg  
2020 2025 2030  
Cys Pro Cys Gly Ala Asn Ile Ser Gly Asn Val Arg Leu Gly Ser Met  
2035 2040 2045  
Arg Ile Thr Gly Pro Lys Thr Cys Met Asn Thr Trp Gln Gly Thr Phe  
2050 2055 2060  
Pro Ile Asn Cys Tyr Thr Glu Gly Gln Cys Ala Pro Lys Pro Pro Thr  
2065 2070 2075 2080  
Asn Tyr Lys Thr Ala Ile Trp Arg Val Ala Ala Ser Glu Tyr Ala Glu

2085	2090	2095
Val Thr Gln His Gly Ser Tyr Ser Tyr Val Thr Gly Leu Thr Thr Asp		
2100	2105	2110
Asn Leu Lys Ile Pro Cys Gln Leu Pro Ser Pro Glu Phe Phe Ser Trp		
2115	2120	2125
Val Asp Gly Val Gln Ile His Arg Phe Ala Pro Thr Pro Lys Pro Phe		
2130	2135	2140
Phe Arg Asp Glu Val Ser Phe Cys Val Gly Leu Asn Ser Tyr Ala Val		
2145	2150	2155
Gly Ser Gln Leu Pro Cys Glu Pro Glu Pro Asp Ala Asp Val Leu Arg		
2165	2170	2175
Ser Met Leu Thr Asp Pro Pro His Ile Thr Ala Glu Thr Ala Ala Arg		
2180	2185	2190
Arg Leu Ala Arg Gly Ser Pro Pro Ser Glu Ala Ser Ser Ser Val Ser		
2195	2200	2205
Gln Leu Ser Ala Pro Ser Leu Arg Ala Thr Cys Thr Thr His Ser Asn		
2210	2215	2220
Thr Tyr Asp Val Asp Met Val Asp Ala Asn Leu Leu Met Glu Gly Gly		
2225	2230	2235
Val Ala Gln Thr Glu Pro Glu Ser Arg Val Pro Val Leu Asp Phe Leu		
2245	2250	2255
Glu Pro Met Ala Glu Glu Glu Ser Asp Leu Glu Pro Ser Ile Pro Ser		
2260	2265	2270
Glu Cys Met Leu Pro Arg Ser Gly Phe Pro Arg Ala Leu Pro Ala Trp		
2275	2280	2285
Ala Arg Pro Asp Tyr Asn Pro Pro Leu Val Glu Ser Trp Arg Arg Pro		
2290	2295	2300
Asp Tyr Gln Pro Pro Thr Val Ala Gly Cys Ala Leu Pro Pro Pro Lys		
2305	2310	2315
		2320

Lys Ala Pro Thr Pro Pro Pro Arg Arg Arg Arg Thr Val Gly Leu Ser  
2325 2330 2335

Glu Ser Thr Ile Ser Glu Ala Leu Gln Gln Leu Ala Ile Lys Thr Phe  
2340 2345 2350

Gly Gln Pro Pro Ser Ser Gly Asp Ala Gly Ser Ser Thr Gly Ala Gly  
2355 2360 2365

Ala Ala Glu Ser Gly Gly Pro Thr Ser Pro Gly Glu Pro Ala Pro Ser  
2370 2375 2380

Glu Thr Gly Ser Ala Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly  
2385 2390 2395 2400

Asp Pro Asp Leu Glu Ser Asp Gln Val Glu Leu Gln Pro Pro Pro Gln  
2405 2410 2415

Gly Gly Gly Val Ala Pro Gly Ser Gly Ser Gly Ser Trp Ser Thr Cys  
2420 2425 2430

Ser Glu Glu Asp Asp Thr Thr Val Cys Cys Ser Met Ser Tyr Ser Trp  
2435 2440 2445

Thr Gly Ala Leu Ile Thr Pro Cys Ser Pro Glu Glu Glu Lys Leu Pro  
2450 2455 2460

Ile Asn Pro Leu Ser Asn Ser Leu Leu Arg Tyr His Asn Lys Val Tyr  
2465 2470 2475 2480

Cys Thr Thr Ser Lys Ser Ala Ser Gln Arg Ala Lys Lys Val Thr Phe  
2485 2490 2495

Asp Arg Thr Gln Val Leu Asp Ala His Tyr Asp Ser Val Leu Lys Asp  
2500 2505 2510

Ile Lys Leu Ala Ala Ser Lys Val Ser Ala Arg Leu Leu Thr Leu Glu  
2515 2520 2525

Glu Ala Cys Gln Leu Thr Pro Pro His Ser Ala Arg Ser Lys Tyr Gly  
2530 2535 2540

Phe Gly Ala Lys Glu Val Arg Ser Leu Ser Gly Arg Ala Val Asn His

2545                      2550                      2555                      2560  
Ile Lys Ser Val Trp Lys Asp Leu Leu Glu Asp Pro Gln Thr Pro Ile  
                         2565                      2570                      2575  
Pro Thr Thr Ile Met Ala Lys Asn Glu Val Phe Cys Val Asp Pro Ala  
                         2580                      2585                      2590  
Lys Gly Gly Lys Lys Pro Ala Arg Leu Ile Val Tyr Pro Asp Leu Gly  
                         2595                      2600                      2605  
Val Arg Val Cys Glu Lys Met Ala Leu Tyr Asp Ile Thr Gln Lys Leu  
                         2610                      2615                      2620  
Pro Gln Ala Val Met Gly Ala Ser Tyr Gly Phe Gln Tyr Ser Pro Ala  
2625                      2630                      2635                      2640  
Gln Arg Val Glu Tyr Leu Leu Lys Ala Trp Ala Glu Lys Lys Asp Pro  
                         2645                      2650                      2655  
Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe Asp Ser Thr Val Thr Glu  
                         2660                      2665                      2670  
Arg Asp Ile Arg Thr Glu Glu Ser Ile Tyr Gln Ala Cys Ser Leu Pro  
                         2675                      2680                      2685  
Glu Glu Ala Arg Thr Ala Ile His Ser Leu Thr Glu Arg Leu Tyr Val  
                         2690                      2695                      2700  
Gly Gly Pro Met Phe Asn Ser Lys Gly Gln Thr Cys Gly Tyr Arg Arg  
2705                      2710                      2715                      2720  
Cys Arg Ala Ser Gly Val Leu Thr Thr Ser Met Gly Asn Thr Ile Thr  
                         2725                      2730                      2735  
Cys Tyr Val Lys Ala Leu Ala Ala Cys Lys Ala Ala Gly Ile Val Ala  
                         2740                      2745                      2750  
Pro Thr Met Leu Val Cys Gly Asp Asp Leu Val Val Ile Ser Glu Ser  
                         2755                      2760                      2765  
Gln Gly Thr Glu Glu Asp Glu Arg Asn Leu Arg Ala Phe Thr Glu Ala  
                         2770                      2775                      2780



Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp Pro Pro Arg Pro Glu Tyr  
 2785                      2790                      2795                      2800  
 Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser Asn Val Ser Val Ala Leu  
                          2805                      2810                      2815  
 Gly Pro Arg Gly Arg Arg Arg Tyr Tyr Leu Thr Arg Asp Pro Thr Thr  
                          2820                      2825                      2830  
 Pro Leu Ala Arg Ala Ala Trp Glu Thr Val Arg His Ser Pro Ile Asn  
                          2835                      2840                      2845  
 Ser Trp Leu Gly Asn Ile Ile Gln Tyr Ala Pro Thr Ile Trp Val Arg  
                          2850                      2855                      2860  
 Met Val Leu Met Thr His Phe Phe Ser Ile Leu Met Val Gln Asp Thr  
 2865                      2870                      2875                      2880  
 Leu Asp Gln Asn Leu Asn Phe Glu Met Tyr Gly Ser Val Tyr Ser Val  
                          2885                      2890                      2895  
 Asn Pro Leu Asp Leu Pro Ala Ile Ile Glu Arg Leu His Gly Leu Asp  
                          2900                      2905                      2910  
 Ala Phe Ser Met His Thr Tyr Ser His His Glu Leu Thr Arg Val Ala  
                          2915                      2920                      2925  
 Ser Ala Leu Arg Lys Leu Gly Ala Pro Pro Leu Arg Val Trp Lys Ser  
                          2930                      2935                      2940  
 Arg Ala Arg Ala Val Arg Ala Ser Leu Ile Ser Arg Gly Gly Lys Ala  
 2945                      2950                      2955                      2960  
 Ala Val Cys Gly Arg Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu  
                          2965                      2970                      2975  
 Lys Leu Thr Pro Leu Pro Glu Ala Arg Leu Leu Asp Leu Ser Ser Trp  
                          2980                      2985                      2990  
 Phe Thr Val Gly Ala Gly Gly Gly Asp Ile Phe His Ser Val Ser Arg  
                          2995                      3000                      3005  
 Ala Arg Pro Arg Ser Leu Leu Phe Gly Leu Leu Leu Leu Phe Val Gly

3010                      3015                      3020  
Val Gly Leu Phe Leu Leu Pro Ala Arg  
3025                      3030

<210> 5

<211> 9674

<212> DNA

<213> Hepatitis C virus

<220>

<221> CDS

<222> (341)..(9442)

<400> 5

acccgccct aatagggcg acactccgcc atgaatcact cccctgtgag gaactactgt 60

cttcacgcag aaagcgtcta gccatggcgt tagtatgagt gtcgtacagc ctccaggccc 120

ccccctcccg ggagagccat agtgggtctgc ggaaccggtg agtacaccgg aattgccggg 180

aagactgggt cctttcttgg ataaaccac tctatgcccg gccatttggg cgtgcccccg 240

caagactgct agccgagtag cgttgggttg cgaaaggcct tgttgtactg cctgataggg 300

tgcttgcgag tgccccggga ggtctcgtag accgtgcacc atg agc aca aat ccc 355

Met Ser Thr Asn Pro

1

5

aaa cct caa aga aaa acc aaa aga aac act aac cgt cgc cca caa gac 403

Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Gln Asp

10

15

20

gtt aag ttt ccg ggc ggc ggc cag atc gtt ggc gga gta tac ttg ttg 451

Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Gly Val Tyr Leu Leu

25

30

35

ccg cgc agg ggc ccc agg ttg ggt gtg cgc gcg aca agg aag gct tcg 499

Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala Thr Arg Lys Ala Ser

40

45

50

gag cgg tcc cag cca cgt ggg agg cgc cag ccc atc ccc aaa cat cgg 547

Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro Ile Pro Lys His Arg

55

60

65

cgc tcc act ggc aag tcc tgg ggg aag cca gga tac ccc tgg ccc ctg 595

Arg Ser Thr Gly Lys Ser Trp Gly Lys Pro Gly Tyr Pro Trp Pro Leu

70

75

80

85

tat ggg aat gag ggc ctc ggt tgg gca gga tgg ctc ctg tcc cct cga 643

Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp Leu Leu Ser Pro Arg

90

95

100

ggt tcc cgt ccc tca tgg ggc ccc aat gac ccc cgg cat agg tcg cgc 691

Gly Ser Arg Pro Ser Trp Gly Pro Asn Asp Pro Arg His Arg Ser Arg

105

110

115

aat gtg ggt aag gtc atc gat acc cta acg tgc ggc ttt gcc gac ctc 739

Asn Val Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu

120

125

130

ttg ggg tac gtc ccc gtc gta ggc gcc ccg ctt agt ggc gtt gcc agt 787

Leu Gly Tyr Val Pro Val Val Gly Ala Pro Leu Ser Gly Val Ala Ser

135

140

145

gct ctc gcg cac ggc gtg aga gtc ctg gag gac ggg gtt aat ttt gca 835

Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Phe Ala

150

155

160

165

aca ggg aac tta cct ggt tgc tcc ttt tct atc ttc ttg ctg gcc cta 883

Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu

170

175

180

ctg tcc tgc atc act act ccg gtc tct gct gtc caa gtg aag aac acc 931

Leu Ser Cys Ile Thr Thr Pro Val Ser Ala Val Gln Val Lys Asn Thr

185

190

195

agc aac gcc tat atg gcg act aac gac tgt tcc aat gac agc atc act 979

Ser Asn Ala Tyr Met Ala Thr Asn Asp Cys Ser Asn Asp Ser Ile Thr

200

205

210

tgg cag ctt gag gcc gca gtc ctc cat gtc ccc ggg tgc gtc ccg tgc 1027

Trp Gln Leu Glu Ala Ala Val Leu His Val Pro Gly Cys Val Pro Cys

215

220

225

gag aaa atg ggg aac aca tca cgg tgc tgg ata cca gtc tca cca aac 1075

Glu Lys Met Gly Asn Thr Ser Arg Cys Trp Ile Pro Val Ser Pro Asn

230

235

240

245

gtg gct gtg cgg cag cct ggc gcc ctc acg cgg ggc ttg cgg acg cac 1123

Val Ala Val Arg Gln Pro Gly Ala Leu Thr Arg Gly Leu Arg Thr His

250

255

260

atc gac atg gtc gtg ttg tcc gcc acg ctc tgc tcc gct ctc tac gtg 1171

Ile Asp Met Val Val Leu Ser Ala Thr Leu Cys Ser Ala Leu Tyr Val

265

270

275

ggg gac ctc tgt ggc ggg gtg atg ctc gcg tcc cag atg ttc att gtc 1219

Gly Asp Leu Cys Gly Gly Val Met Leu Ala Ser Gln Met Phe Ile Val

280

285

290

tcg ccg cag cac cac tgg ttc gtg cag gaa tgc aat tgc tcc atc tac 1267

Ser Pro Gln His His Trp Phe Val Gln Glu Cys Asn Cys Ser Ile Tyr

295

300

305

cct ggc gcc atc act ggg cac cgt atg gca tgg gac atg atg atg aac 1315

Pro Gly Ala Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn

310

315

320

325

tgg tcg ccc acg acc acc atg atc ctg gcg tac gtg atg cgc gtt ccc 1363

Trp Ser Pro Thr Thr Thr Met Ile Leu Ala Tyr Val Met Arg Val Pro

330

335

340

gag gtc atc ata gac atc att agc gga gct cac tgg ggc gtc atg ttt 1411

Glu Val Ile Ile Asp Ile Ile Ser Gly Ala His Trp Gly Val Met Phe

345

350

355

ggc ctg gcc tac ttc tct atg cag gga gcg tgg gcg aag gtc gtt gtc 1459

Gly Leu Ala Tyr Phe Ser Met Gln Gly Ala Trp Ala Lys Val Val Val

360

365

370

atc ctc ctg ctg gcc tct ggg gtg gac gcg tac acc acc acg act ggg 1507

Ile Leu Leu Leu Ala Ser Gly Val Asp Ala Tyr Thr Thr Thr Thr Gly

375

380

385

agc gct gct ggg cgc act acc agt agc ctg gcc agc gcc ttc tcc cct 1555

Ser Ala Ala Gly Arg Thr Thr Ser Ser Leu Ala Ser Ala Phe Ser Pro

390

395

400

405

ggc gct cgg cag aac att cag ctc att aat acc aat ggt agc tgg cac 1603

Gly Ala Arg Gln Asn Ile Gln Leu Ile Asn Thr Asn Gly Ser Trp His

410

415

420

atc aac cgc acc gcc ctg aat tgc aac gat tcc ttg cac acc ggc ttc 1651

Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser Leu His Thr Gly Phe

425

430

435

ttc acg gcc ctg ttc tac atc cat aag ttc aac tcg tcg gga tgt ccc 1699

Phe Thr Ala Leu Phe Tyr Ile His Lys Phe Asn Ser Ser Gly Cys Pro

440

445

450

gag cgc ctg tcc gcc tgt cgc aac atc gag gac ttc cgg ata gga tgg 1747

Glu Arg Leu Ser Ala Cys Arg Asn Ile Glu Asp Phe Arg Ile Gly Trp

455

460

465

ggc gcc ctg caa tac gac gac aat gtc acc aat cca gaa gat atg agg 1795  
 Gly Ala Leu Gln Tyr Asp Asp Asn Val Thr Asn Pro Glu Asp Met Arg  
 470 475 480 485

cca tat tgc tgg cac tac cca cca aaa cag tgt ggc gta gtc ccc gca 1843  
 Pro Tyr Cys Trp His Tyr Pro Pro Lys Gln Cys Gly Val Val Pro Ala  
 490 495 500

ggg acc gtg tgc ggc cca gtg tac tgt ttc acc cct agc ccg gtg gta 1891  
 Gly Thr Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val Val  
 505 510 515

gtg ggc acg acc gat aga ctt gga gtg cct act tac acg tgg gga gag 1939  
 Val Gly Thr Thr Asp Arg Leu Gly Val Pro Thr Tyr Thr Trp Gly Glu  
 520 525 530

aat gag aca gat gtc ttc cta ttg aac agc acc cga cca ccg tcg ggg 1987  
 Asn Glu Thr Asp Val Phe Leu Leu Asn Ser Thr Arg Pro Pro Ser Gly  
 535 540 545

tca tgg ttt ggc tgc acg tgg atg aac tcc act ggc ttc acc aag acc 2035  
 Ser Trp Phe Gly Cys Thr Trp Met Asn Ser Thr Gly Phe Thr Lys Thr  
 550 555 560 565

tgc ggc gca cca ccc tgc cgc act aga gct gac ttc aat acc agc aca 2083  
 Cys Gly Ala Pro Pro Cys Arg Thr Arg Ala Asp Phe Asn Thr Ser Thr  
 570 575 580

gat ctg ttg tgc ccc acg gac tgt ttt aga aaa cat cct gaa gcc act 2131

Asp Leu Leu Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala Thr  
 585 590 595

tac atc aaa tgt ggt tcc ggg cct tgg ctc acg cca aag tgt ctg gtt 2179  
 Tyr Ile Lys Cys Gly Ser Gly Pro Trp Leu Thr Pro Lys Cys Leu Val  
 600 605 610

gac tac ccc tac agg ctc tgg cat tac cct tgc aca gtc aat tac tcc 2227  
 Asp Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Tyr Ser  
 615 620 625

acc ttc aag atc agg atg tat gtg ggg gga gtt gag cac agg ctc atg 2275  
 Thr Phe Lys Ile Arg Met Tyr Val Gly Gly Val Glu His Arg Leu Met  
 630 635 640 645

gcc gcg tgc aat ttc act cgt ggg gat cgc tgc aac ttg gag gat agg 2323  
 Ala Ala Cys Asn Phe Thr Arg Gly Asp Arg Cys Asn Leu Glu Asp Arg  
 650 655 660

gac aga agt caa cag act cct ctg ttg cac tcc acc acg gaa tgg gcc 2371  
 Asp Arg Ser Gln Gln Thr Pro Leu Leu His Ser Thr Thr Glu Trp Ala  
 665 670 675

att ttg ccc tgc tct ttc tca gac ttg ccc gct ttg tcg act ggt ctt 2419  
 Ile Leu Pro Cys Ser Phe Ser Asp Leu Pro Ala Leu Ser Thr Gly Leu  
 680 685 690

ctc cac ctc cac caa aat atc gtg gac gta caa tat atg tat ggc ctg 2467  
 Leu His Leu His Gln Asn Ile Val Asp Val Gln Tyr Met Tyr Gly Leu



695

700

705

tca cct gcc ctc aca caa tat atc gtt cga tgg gag tgg gta gta ctc 2515  
Ser Pro Ala Leu Thr Gln Tyr Ile Val Arg Trp Glu Trp Val Val Leu  
710 715 720 725

tta ttc ctg ctc cta gcg gac gcc agg gtc tgc gcc tgc ttg tgg atg 2563  
Leu Phe Leu Leu Leu Ala Asp Ala Arg Val Cys Ala Cys Leu Trp Met  
730 735 740

ctc atc ttg ctg ggc caa gcc gaa gca gca ctg gag aag ctg gtc gtc 2611  
Leu Ile Leu Leu Gly Gln Ala Glu Ala Ala Leu Glu Lys Leu Val Val  
745 750 755

ttg cac gct gcg agc gca gct agc tgc aat ggc ttc ctg tat ttt gtc 2659  
Leu His Ala Ala Ser Ala Ala Ser Cys Asn Gly Phe Leu Tyr Phe Val  
760 765 770

atc ttt ctc gtg gct gct tgg cac atc aag ggt agg gtg gtc ccc ttg 2707  
Ile Phe Leu Val Ala Ala Trp His Ile Lys Gly Arg Val Val Pro Leu  
775 780 785

gct gct tat tcc ctt act ggc ctg tgg ccg ttc tgc cta ctg ctc cta 2755  
Ala Ala Tyr Ser Leu Thr Gly Leu Trp Pro Phe Cys Leu Leu Leu Leu  
790 795 800 805

gca ctg ccc cag cag gct tac gcc tat gat gca tct gtg cac gga cag 2803  
Ala Leu Pro Gln Gln Ala Tyr Ala Tyr Asp Ala Ser Val His Gly Gln  
810 815 820

gtg ggc gcg gct ttg cta gta ctg att acc ctc ttt aca ctc acc ccg 2851

Val Gly Ala Ala Leu Leu Val Leu Ile Thr Leu Phe Thr Leu Thr Pro

825

830

835

ggg tat aag acc ctt ctc agc cag tcc ctg tgg tgg ttg tgc tat ctc 2899

Gly Tyr Lys Thr Leu Leu Ser Gln Ser Leu Trp Trp Leu Cys Tyr Leu

840

845

850

ctg acc ctg gcg gaa acc atg gtc cag gag tgg gca cca tcc atg cag 2947

Leu Thr Leu Ala Glu Thr Met Val Gln Glu Trp Ala Pro Ser Met Gln

855

860

865

gcg cgc ggc ggc cgt gat ggc atc ata tgg gcc gcc acc ata ttt tgc 2995

Ala Arg Gly Gly Arg Asp Gly Ile Ile Trp Ala Ala Thr Ile Phe Cys

870

875

880

885

ccg ggc gta gtg ttt gac ata acc aag tgg ctc tta gcg gtg ctt ggg 3043

Pro Gly Val Val Phe Asp Ile Thr Lys Trp Leu Leu Ala Val Leu Gly

890

895

900

cct ggt tac ctc cta aga ggt gct ttg acg cgc gtg cca tat ttc gtc 3091

Pro Gly Tyr Leu Leu Arg Gly Ala Leu Thr Arg Val Pro Tyr Phe Val

905

910

915

aga gcc cac gct ctg ctg aga atg tgc act atg gtg agg cac ctc gcg 3139

Arg Ala His Ala Leu Leu Arg Met Cys Thr Met Val Arg His Leu Ala

920

925

930

ggg ggt agg tac gtc cag atg gcg cta tta gcc ctt ggc agg tgg act 3187

Gly Gly Arg Tyr Val Gln Met Ala Leu Leu Ala Leu Gly Arg Trp Thr

935

940

945

ggc act tac atc tat gac cac ctc acc cct atg tcg gat tgg gct gct 3235

Gly Thr Tyr Ile Tyr Asp His Leu Thr Pro Met Ser Asp Trp Ala Ala

950

955

960

965

agc ggc ctg cgg gac ttg gcg gtc gct gtg gag cct atc atc ttc agt 3283

Ser Gly Leu Arg Asp Leu Ala Val Ala Val Glu Pro Ile Ile Phe Ser

970

975

980

ccg atg gag aag aaa gtc atc gtt tgg gga gcg gag acg gct gcg tgc 3331

Pro Met Glu Lys Lys Val Ile Val Trp Gly Ala Glu Thr Ala Ala Cys

985

990

995

ggg gac atc ttg cac gga ctt ccc gtg tcc gcc cga ctc ggt cgg gag 3379

Gly Asp Ile Leu His Gly Leu Pro Val Ser Ala Arg Leu Gly Arg Glu

1000

1005

1010

atc ctc ctt ggc cca gct gat ggc tac acc tcc aag ggg tgg aag ctt 3427

Ile Leu Leu Gly Pro Ala Asp Gly Tyr Thr Ser Lys Gly Trp Lys Leu

1015

1020

1025

ctc gcc ccc atc acc gct tac gcc cag cag aca cga ggt ctc ttg ggc 3475

Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu Leu Gly

1030

1035

1040

1045

tct ata gtg gtg agc atg acg ggg cgt gac aag aca gaa cag gcc ggg 3523

Ser Ile Val Val Ser Met Thr Gly Arg Asp Lys Thr Glu Gln Ala Gly  
1050 1055 1060

gag gtc caa gtc ctg tcc aca gtc act cag tcc ttc ctc gga aca tcc 3571  
Glu Val Gln Val Leu Ser Thr Val Thr Gln Ser Phe Leu Gly Thr Ser  
1065 1070 1075

att tcg ggg gtc tta tgg act gtt tac cac gga gct ggc aac aag aca 3619  
Ile Ser Gly Val Leu Trp Thr Val Tyr His Gly Ala Gly Asn Lys Thr  
1080 1085 1090

cta gcc ggc tcg cgg ggc ccg gtc acg cag atg tac tcg agc gcc gag 3667  
Leu Ala Gly Ser Arg Gly Pro Val Thr Gln Met Tyr Ser Ser Ala Glu  
1095 1100 1105

ggg gac ttg gtc ggg tgg ccc agc cct cct ggg acc aaa tct ttg gag 3715  
Gly Asp Leu Val Gly Trp Pro Ser Pro Pro Gly Thr Lys Ser Leu Glu  
1110 1115 1120 1125

ccg tgt acg tgt gga gcg gtc gac ctg tat ttg gtc acg cgg aac gct 3763  
Pro Cys Thr Cys Gly Ala Val Asp Leu Tyr Leu Val Thr Arg Asn Ala  
1130 1135 1140

gat gtc atc ccg gct cga aga cgc ggg gac aag cgg gga gcg ctg ctc 3811  
Asp Val Ile Pro Ala Arg Arg Arg Gly Asp Lys Arg Gly Ala Leu Leu  
1145 1150 1155

tcc ccg aga ccc ctt tcg acc ttg aag ggg tcc tcg ggg gga cct gtg 3859  
Ser Pro Arg Pro Leu Ser Thr Leu Lys Gly Ser Ser Gly Gly Pro Val

1160 1165 1170

ctt tgc cct agg ggc cac gct gtc gga atc ttc cgg gca gct gtg tgc 3907  
Leu Cys Pro Arg Gly His Ala Val Gly Ile Phe Arg Ala Ala Val Cys  
1175 1180 1185

tct cgg ggt gtg gct aag tcc ata gat ttc atc ccc gtt gag acg ctc 3955  
Ser Arg Gly Val Ala Lys Ser Ile Asp Phe Ile Pro Val Glu Thr Leu  
1190 1195 1200 1205

gac atc gtc acg cgg tct ccc acc ttt agt gac aac agc aca cca cca 4003  
Asp Ile Val Thr Arg Ser Pro Thr Phe Ser Asp Asn Ser Thr Pro Pro  
1210 1215 1220

gct gtg ccc cag acc tat cag gtg ggg tac ttg cac gcc ccc act ggc 4051  
Ala Val Pro Gln Thr Tyr Gln Val Gly Tyr Leu His Ala Pro Thr Gly  
1225 1230 1235

agt gga aaa agc acc aag gtc ccc gtc gcg tac gcc gcc cag ggg tat 4099  
Ser Gly Lys Ser Thr Lys Val Pro Val Ala Tyr Ala Ala Gln Gly Tyr  
1240 1245 1250

aaa gtg ctg gtg ctc aat ccc tcg gtg gct gcc acc ctg gga ttt ggg 4147  
Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly  
1255 1260 1265

gcg tac ttg tcc aag gca cat ggc atc aac ccc aac att agg act gga 4195  
Ala Tyr Leu Ser Lys Ala His Gly Ile Asn Pro Asn Ile Arg Thr Gly  
1270 1275 1280 1285

gtc aga act gtg acg acc ggg gag ccc att aca tac tcc acg tat ggt 4243

Val Arg Thr Val Thr Thr Gly Glu Pro Ile Thr Tyr Ser Thr Tyr Gly

1290

1295

1300

aaa ttc ctc gcc gat ggg ggc tgc gca ggc ggc gcc tat gac atc atc 4291

Lys Phe Leu Ala Asp Gly Gly Cys Ala Gly Gly Ala Tyr Asp Ile Ile

1305

1310

1315

ata tgc gat gaa tgc cac tct gtg gat gct acc act att ctc ggc atc 4339

Ile Cys Asp Glu Cys His Ser Val Asp Ala Thr Thr Ile Leu Gly Ile

1320

1325

1330

ggg aca gtc ctt gac caa gca gag aca gcc ggg gtc agg cta act gta 4387

Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Val Arg Leu Thr Val

1335

1340

1345

ctg gcc acg gcc acg ccc ccc ggg tcg gtg aca acc ccc cat ccc aat 4435

Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Thr Pro His Pro Asn

1350

1355

1360

1365

ata gag gag gta gcc ctc gga cag gag ggt gag atc ccc ttc tat ggg 4483

Ile Glu Glu Val Ala Leu Gly Gln Glu Gly Glu Ile Pro Phe Tyr Gly

1370

1375

1380

agg gcg ttt ccc ctg tct tac atc aag gga ggg agg cac ttg att ttc 4531

Arg Ala Phe Pro Leu Ser Tyr Ile Lys Gly Gly Arg His Leu Ile Phe

1385

1390

1395

tgc cac tca aag aaa aag tgt gac gag ctc gca acg gcc ctt cgg ggc 4579

Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Thr Ala Leu Arg Gly

1400

1405

1410

atg ggc ttg aac gct gtg gca tat tac aga ggg ttg gac gtc tcc ata 4627

Met Gly Leu Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Ile

1415

1420

1425

ata cca act caa gga gat gtg gtg gtc gtt gcc acc gac gcc ctc atg 4675

Ile Pro Thr Gln Gly Asp Val Val Val Val Ala Thr Asp Ala Leu Met

1430

1435

1440

1445

acg ggg tat act gga gac ttt gac tcc gtg atc gac tgc aac gta gcg 4723

Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Val Ala

1450

1455

1460

gtc acc cag gcc gta gac ttc agc ctg gac ccc acc ttc act ata acc 4771

Val Thr Gln Ala Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Thr

1465

1470

1475

aca cag act gtc ccg caa gac gct gtc tca cgt agt cag cgc cga ggg 4819

Thr Gln Thr Val Pro Gln Asp Ala Val Ser Arg Ser Gln Arg Arg Gly

1480

1485

1490

cgc acg ggt aga gga aga ctg ggc att tat agg tat gtt tcc act ggt 4867

Arg Thr Gly Arg Gly Arg Leu Gly Ile Tyr Arg Tyr Val Ser Thr Gly

1495

1500

1505

gag cga gcc tca gga atg ttt gac agt gta gta ctc tgt gag tgc tac 4915

Glu Arg Ala Ser Gly Met Phe Asp Ser Val Val Leu Cys Glu Cys Tyr  
1510 1515 1520 1525

gac gca gga gct gct tgg tat gag ctc tca cca gtg gag acg acc gtc 4963  
Asp Ala Gly Ala Ala Trp Tyr Glu Leu Ser Pro Val Glu Thr Thr Val  
1530 1535 1540

agg ctc agg gcg tat ttc aac acg cct ggc ttg cct gtg tgc cag gac 5011  
Arg Leu Arg Ala Tyr Phe Asn Thr Pro Gly Leu Pro Val Cys Gln Asp  
1545 1550 1555

cac ctt gag ttt tgg gag gca gtt ttc acc ggc ctc aca cac ata gac 5059  
His Leu Glu Phe Trp Glu Ala Val Phe Thr Gly Leu Thr His Ile Asp  
1560 1565 1570

gct cat ttc ctt tcc cag aca aag cag tcg ggg gaa aat ttc gca tac 5107  
Ala His Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Phe Ala Tyr  
1575 1580 1585

tta gta gcc tat cag gcc aca gtg tgc gcc agg gcc aaa gcg ccc ccc 5155  
Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Lys Ala Pro Pro  
1590 1595 1600 1605

ccg tcc tgg gac gtc atg tgg aag tgc ttg act cga ctc aag ccc acg 5203  
Pro Ser Trp Asp Val Met Trp Lys Cys Leu Thr Arg Leu Lys Pro Thr  
1610 1615 1620

ctt gtg ggc cct aca cct ctc ctg tac cgt ttg ggc tct gtt acc aac 5251  
Leu Val Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ser Val Thr Asn



1625	1630	1635	
gag gtc acc ctt aca cac ccc gtg aca aaa tac atc gcc aca tgc atg 5299			
Glu Val Thr Leu Thr His Pro Val Thr Lys Tyr Ile Ala Thr Cys Met			
1640	1645	1650	
caa gct gac ctc gag gtc atg acc agc acg tgg gtc ctg gct ggg gga 5347			
Gln Ala Asp Leu Glu Val Met Thr Ser Thr Trp Val Leu Ala Gly Gly			
1655	1660	1665	
gtc tta gca gcc gtc gcc gcg tat tgc tta gcg acc ggg tgt gtt tcc 5395			
Val Leu Ala Ala Val Ala Ala Tyr Cys Leu Ala Thr Gly Cys Val Ser			
1670	1675	1680	1685
atc att ggc cgt tta cac atc aac cag cga gct gtc gtc gct ccg gac 5443			
Ile Ile Gly Arg Leu His Ile Asn Gln Arg Ala Val Val Ala Pro Asp			
1690	1695	1700	
aag gag gtc ctc tat gag gct ttt gat gag atg gag gaa tgt gcc tcc 5491			
Lys Glu Val Leu Tyr Glu Ala Phe Asp Glu Met Glu Glu Cys Ala Ser			
1705	1710	1715	
aga gcg gct ctc ctt gaa gag ggg cag cgg ata gcc gag atg ctg aag 5539			
Arg Ala Ala Leu Leu Glu Glu Gly Gln Arg Ile Ala Glu Met Leu Lys			
1720	1725	1730	
tcc aag atc caa ggc tta ttg cag caa gcc tct aaa cag gcc cag gac 5587			
Ser Lys Ile Gln Gly Leu Leu Gln Gln Ala Ser Lys Gln Ala Gln Asp			
1735	1740	1745	

ata caa ccc gct gtg caa gct tcg tgg ccc aag atg gag caa ttc tgg 5635

Ile Gln Pro Ala Val Gln Ala Ser Trp Pro Lys Met Glu Gln Phe Trp

1750 1755 1760 1765

gcc aaa cat atg tgg aac ttc ata agc ggc att cag tac ctc gca gga 5683

Ala Lys His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala Gly

1770 1775 1780

ctg tca aca ctg cca ggg aac cct gct gtg gct tcc atg atg gca ttc 5731

Leu Ser Thr Leu Pro Gly Asn Pro Ala Val Ala Ser Met Met Ala Phe

1785 1790 1795

agc gcc gcc ctc acc agt ccg ttg tca act agc acc acc atc ctt ctt 5779

Ser Ala Ala Leu Thr Ser Pro Leu Ser Thr Ser Thr Thr Ile Leu Leu

1800 1805 1810

aac att ctg ggg ggc tgg ctg gcg tcc caa att gcg cca ccc gcg ggg 5827

Asn Ile Leu Gly Gly Trp Leu Ala Ser Gln Ile Ala Pro Pro Ala Gly

1815 1820 1825

gcc act ggc ttt gtt gtc agt ggc ctg gtg gga gct gct gtt ggc agc 5875

Ala Thr Gly Phe Val Val Ser Gly Leu Val Gly Ala Ala Val Gly Ser

1830 1835 1840 1845

ata ggc ttg ggt aaa gtg ctg gtg gac atc ctg gca ggg tat ggt gcg 5923

Ile Gly Leu Gly Lys Val Leu Val Asp Ile Leu Ala Gly Tyr Gly Ala

1850 1855 1860

ggc att tcg ggg gcc ctc gtc gcg ttt aag atc atg tct ggc gag aag 5971

Gly Ile Ser Gly Ala Leu Val Ala Phe Lys Ile Met Ser Gly Glu Lys

1865

1870

1875

ccc tcc atg gag gat gtc atc aac ttg ctg cct ggg att ctg tct cca 6019

Pro Ser Met Glu Asp Val Ile Asn Leu Leu Pro Gly Ile Leu Ser Pro

1880

1885

1890

ggt gct ctg gtg gtg gga gtc atc tgc gcg gcc att ctg cgc cgc cat 6067

Gly Ala Leu Val Val Gly Val Ile Cys Ala Ala Ile Leu Arg Arg His

1895

1900

1905

gtg gga ccg ggg gaa ggc gcg gtc caa tgg atg aac agg ctt atc gcc 6115

Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile Ala

1910

1915

1920

1925

ttc gct tcc aga gga aac cac gtc gcc cct act cac tac gtg acg gag 6163

Phe Ala Ser Arg Gly Asn His Val Ala Pro Thr His Tyr Val Thr Glu

1930

1935

1940

tcg gat gcg tcg cag cgt gtc acc caa ctg ctt ggc tct ctc act ata 6211

Ser Asp Ala Ser Gln Arg Val Thr Gln Leu Leu Gly Ser Leu Thr Ile

1945

1950

1955

act agt cta ctc agg aga ctt cac aac tgg atc act gag gat tgc ccc 6259

Thr Ser Leu Leu Arg Arg Leu His Asn Trp Ile Thr Glu Asp Cys Pro

1960

1965

1970

atc cca tgc gcc ggc tcg tgg ctc cgc gat gtg tgg gac tgg gtc tgt 6307

Ile Pro Cys Ala Gly Ser Trp Leu Arg Asp Val Trp Asp Trp Val Cys  
 1975 1980 1985

acc atc cta aca gac ttt aag aac tgg ctg acc tcc aag ctg ttc cca 6355  
 Thr Ile Leu Thr Asp Phe Lys Asn Trp Leu Thr Ser Lys Leu Phe Pro  
 1990 1995 2000 2005

aag atg cct ggc ctc ccc ttt atc tct tgc caa aag ggg tac aag ggc 6403  
 Lys Met Pro Gly Leu Pro Phe Ile Ser Cys Gln Lys Gly Tyr Lys Gly  
 2010 2015 2020

gtg tgg gcc ggc act ggc atc atg acc aca cga tgc ccc tgc ggc gcc 6451  
 Val Trp Ala Gly Thr Gly Ile Met Thr Thr Arg Cys Pro Cys Gly Ala  
 2025 2030 2035

aac atc tct ggc aac gtc cgc ttg ggc tct atg aga atc aca gga ccc 6499  
 Asn Ile Ser Gly Asn Val Arg Leu Gly Ser Met Arg Ile Thr Gly Pro  
 2040 2045 2050

aaa acc tgc atg aac acc tgg cag ggg acc ttt cct atc aat tgt tat 6547  
 Lys Thr Cys Met Asn Thr Trp Gln Gly Thr Phe Pro Ile Asn Cys Tyr  
 2055 2060 2065

aca gaa ggc cag tgc ttg ccg aaa ccc gcg tta aac ttc aag acc gcc 6595  
 Thr Glu Gly Gln Cys Leu Pro Lys Pro Ala Leu Asn Phe Lys Thr Ala  
 2070 2075 2080 2085

atc tgg aga gtg gcg gcc tca gag tac gcg gaa gtg acg cag cac gga 6643  
 Ile Trp Arg Val Ala Ala Ser Glu Tyr Ala Glu Val Thr Gln His Gly

2090	2095	2100	
tca tat gcc tat ata aca ggg ctg acc act gac aac tta aaa gtc cct	6691		
Ser Tyr Ala Tyr Ile Thr Gly Leu Thr Thr Asp Asn Leu Lys Val Pro			
2105	2110	2115	
tgc caa ctc ccc tct cca gag ttt ttc tct tgg gtg gac gga gta caa	6739		
Cys Gln Leu Pro Ser Pro Glu Phe Phe Ser Trp Val Asp Gly Val Gln			
2120	2125	2130	
atc cat agg tcc gcc ccc aca cca aag ccg ttt ttc cgg gat gag gtc	6787		
Ile His Arg Ser Ala Pro Thr Pro Lys Pro Phe Phe Arg Asp Glu Val			
2135	2140	2145	
tcg ttc agc gtt ggg ctc aat tca ttt gtc gtc ggg tct cag ctt ccc	6835		
Ser Phe Ser Val Gly Leu Asn Ser Phe Val Val Gly Ser Gln Leu Pro			
2150	2155	2160	2165
tgt gac cct gag ccc gac act gag gta gtg atg tcc atg cta aca gac	6883		
Cys Asp Pro Glu Pro Asp Thr Glu Val Val Met Ser Met Leu Thr Asp			
2170	2175	2180	
cca tcc cat atc acg gcg gag gct gca gcg cgg cgt tta gcg cgg ggg	6931		
Pro Ser His Ile Thr Ala Glu Ala Ala Ala Arg Arg Leu Ala Arg Gly			
2185	2190	2195	
tca ccc cca tct gag gca agc tcc tca gcg agc cag ctg tcg gcg cca	6979		
Ser Pro Pro Ser Glu Ala Ser Ser Ser Ala Ser Gln Leu Ser Ala Pro			
2200	2205	2210	

tcg ctg cga gcc acc tgc acc acc cac ggt agg acc tat gat gtg gac 7027

Ser Leu Arg Ala Thr Cys Thr Thr His Gly Arg Thr Tyr Asp Val Asp

2215

2220

2225

atg gtg gat gcc aac ctg ttc atg ggg ggc ggc gtg att cgg ata gag 7075

Met Val Asp Ala Asn Leu Phe Met Gly Gly Gly Val Ile Arg Ile Glu

2230

2235

2240

2245

tct gag tcc aaa gtg gtc gtt ctg gac tcc ctc gac tca atg acc gag 7123

Ser Glu Ser Lys Val Val Val Leu Asp Ser Leu Asp Ser Met Thr Glu

2250

2255

2260

gaa gag ggc gac ctt gag cct tca gta cca tcg gag tat atg ctc ccc 7171

Glu Glu Gly Asp Leu Glu Pro Ser Val Pro Ser Glu Tyr Met Leu Pro

2265

2270

2275

agg aag agg ttc cca ccg gcc tta ccg gct tgg gcg cgg cct gat tac 7219

Arg Lys Arg Phe Pro Pro Ala Leu Pro Ala Trp Ala Arg Pro Asp Tyr

2280

2285

2290

aac cca ccg ctt gtg gaa tcg tgg aag agg cca gat tac caa cca ccc 7267

Asn Pro Pro Leu Val Glu Ser Trp Lys Arg Pro Asp Tyr Gln Pro Pro

2295

2300

2305

act gtt gcg ggc tgt gct ctc ccc ccc ccc aaa aag acc ccg acg cct 7315

Thr Val Ala Gly Cys Ala Leu Pro Pro Pro Lys Lys Thr Pro Thr Pro

2310

2315

2320

2325

cct cca agg aga cgc cgg aca gtg ggt ctg agc gag agc acc ata gga 7363  
Pro Pro Arg Arg Arg Arg Thr Val Gly Leu Ser Glu Ser Thr Ile Gly

2330

2335

2340

gat gcc ctc caa cag ctg gcc atc aag tcc ttt ggc cag ccc ccc cca 7411  
Asp Ala Leu Gln Gln Leu Ala Ile Lys Ser Phe Gly Gln Pro Pro Pro

2345

2350

2355

agc ggc gat tca ggc ctt tcc acg ggg gcg gac gcc gcc gac tcc ggc 7459  
Ser Gly Asp Ser Gly Leu Ser Thr Gly Ala Asp Ala Ala Asp Ser Gly

2360

2365

2370

gat cgg aca ccc cct gac gag ttg gct ctt tcg gag aca ggt tct acc 7507  
Asp Arg Thr Pro Pro Asp Glu Leu Ala Leu Ser Glu Thr Gly Ser Thr

2375

2380

2385

tcc tcc atg ccc ccc ctc gag ggg gag cct ggg gac cca gac ctg gag 7555  
Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu Glu  
2390 2395 2400 2405

cct gag cag gta gag ctt caa cct cct ccc cag ggg ggg gag gca gct 7603  
Pro Glu Gln Val Glu Leu Gln Pro Pro Pro Gln Gly Gly Glu Ala Ala

2410

2415

2420

ccc ggc tcg gac tcg ggg tcc tgg tct act tgc tcc gag gag gat gac 7651  
Pro Gly Ser Asp Ser Gly Ser Trp Ser Thr Cys Ser Glu Glu Asp Asp

2425

2430

2435

tcc gtc gtg tgc tgc tcc atg tca tat tcc tgg acc ggg gct cta ata 7699

Ser Val Val Cys Cys Ser Met Ser Tyr Ser Trp Thr Gly Ala Leu Ile  
2440 2445 2450

act cct tgt agc ccc gaa gag gaa aag ttg cca att aac tcc ttg agc 7747  
Thr Pro Cys Ser Pro Glu Glu Glu Lys Leu Pro Ile Asn Ser Leu Ser  
2455 2460 2465

aac tcg ctg ttg cga tac cat aac aag gta tac tgt act aca tca aag 7795  
Asn Ser Leu Leu Arg Tyr His Asn Lys Val Tyr Cys Thr Thr Ser Lys  
2470 2475 2480 2485

agt gcc tca cta agg gct aaa aag gta act ttt gat agg atg caa gtg 7843  
Ser Ala Ser Leu Arg Ala Lys Lys Val Thr Phe Asp Arg Met Gln Val  
2490 2495 2500

ctc gac gcc tat tat gat tca gtc tta aag gac atc aag cta gcg gcc 7891  
Leu Asp Ala Tyr Tyr Asp Ser Val Leu Lys Asp Ile Lys Leu Ala Ala  
2505 2510 2515

tcc aag gtc agc gca agg ctc ctc acc tta gag gag gcg tgc caa ttg 7939  
Ser Lys Val Ser Ala Arg Leu Leu Thr Leu Glu Glu Ala Cys Gln Leu  
2520 2525 2530

acc cca ccc cac tct gca aga tcc aag tat ggg ttt ggg gct aag gag 7987  
Thr Pro Pro His Ser Ala Arg Ser Lys Tyr Gly Phe Gly Ala Lys Glu  
2535 2540 2545

gtc cgc agc ttg tcc ggg agg gcc gtc aac cac atc aag tcc gtg tgg 8035  
Val Arg Ser Leu Ser Gly Arg Ala Val Asn His Ile Lys Ser Val Trp



2550                      2555                      2560                      2565

aag gac ctc ttg gaa gac tca caa aca cca att cct aca acc atc atg 8083  
Lys Asp Leu Leu Glu Asp Ser Gln Thr Pro Ile Pro Thr Thr Ile Met  
2570                      2575                      2580

gcc aaa aat gag gtg ttc tgc gtg gac ccc gcc aag ggg ggt aaa aaa 8131  
Ala Lys Asn Glu Val Phe Cys Val Asp Pro Ala Lys Gly Gly Lys Lys  
2585                      2590                      2595

cca gct cgc ctt atc gtt tac cct gac ctc ggc gtc agg gtc tgc gag 8179  
Pro Ala Arg Leu Ile Val Tyr Pro Asp Leu Gly Val Arg Val Cys Glu  
2600                      2605                      2610

aag atg gcc ctt tat gat gtc aca caa aag ctt cct cag gcg gtg atg 8227  
Lys Met Ala Leu Tyr Asp Val Thr Gln Lys Leu Pro Gln Ala Val Met  
2615                      2620                      2625

ggg gct tct tat ggc ttc cag tac tcc ccc gct cag cgg gtg gag ttt 8275  
Gly Ala Ser Tyr Gly Phe Gln Tyr Ser Pro Ala Gln Arg Val Glu Phe  
2630                      2635                      2640                      2645

ctc ttg aag gca tgg gcg gaa aag aga gac cct atg ggt ttt tcg tat 8323  
Leu Leu Lys Ala Trp Ala Glu Lys Arg Asp Pro Met Gly Phe Ser Tyr  
2650                      2655                      2660

gat acc cga tgc ttt gac tca acc gtc act gag aga gac atc agg act 8371  
Asp Thr Arg Cys Phe Asp Ser Thr Val Thr Glu Arg Asp Ile Arg Thr  
2665                      2670                      2675

gag gag tcc ata tac cag gcc tgc tcc tta ccc gag gag gcc cga act 8419

Glu Glu Ser Ile Tyr Gln Ala Cys Ser Leu Pro Glu Glu Ala Arg Thr

2680

2685

2690

gcc ata cac tcg ctg act gag aga ctc tat gtg gga ggg ccc atg ttc 8467

Ala Ile His Ser Leu Thr Glu Arg Leu Tyr Val Gly Gly Pro Met Phe

2695

2700

2705

aac agc aag ggc cag tcc tgc ggg tac agg cgt tgc cgc gcc agc ggg 8515

Asn Ser Lys Gly Gln Ser Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly

2710

2715

2720

2725

gtg ctt acc act agt atg ggg aac acc atc aca tgc tat gta aaa gcc 8563

Val Leu Thr Thr Ser Met Gly Asn Thr Ile Thr Cys Tyr Val Lys Ala

2730

2735

2740

cta gcg gct tgc aag gct gcg ggg ata att gcg ccc acg atg ctg gta 8611

Leu Ala Ala Cys Lys Ala Ala Gly Ile Ile Ala Pro Thr Met Leu Val

2745

2750

2755

tgc ggc gac gac ttg gtc gtc atc tca gaa agc cag ggg act gag gag 8659

Cys Gly Asp Asp Leu Val Val Ile Ser Glu Ser Gln Gly Thr Glu Glu

2760

2765

2770

gac gag cgg aac ctg aga gcc ttc acg gag gct atg acc agg tat tct 8707

Asp Glu Arg Asn Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser

2775

2780

2785

gcc cct cct ggt gac ccc ccc aga ccg gaa tat gac ctg gag cta ata 8755  
Ala Pro Pro Gly Asp Pro Pro Arg Pro Glu Tyr Asp Leu Glu Leu Ile  
2790 2795 2800 2805

aca tct tgt tcc tca aac gtg tct gtg gca ctt ggc cca cag ggc cgc 8803  
Thr Ser Cys Ser Ser Asn Val Ser Val Ala Leu Gly Pro Gln Gly Arg  
2810 2815 2820

cgc aga tac tac ctg acc aga gac ccc acc act tca att gcc cgg gct 8851  
Arg Arg Tyr Tyr Leu Thr Arg Asp Pro Thr Thr Ser Ile Ala Arg Ala  
2825 2830 2835

gcc tgg gaa aca gtt aga cac tcc cct gtc aat tca tgg ctg gga aac 8899  
Ala Trp Glu Thr Val Arg His Ser Pro Val Asn Ser Trp Leu Gly Asn  
2840 2845 2850

atc atc cag tac gct cca acc ata tgg gtt cgc atg gtc ctg atg aca 8947  
Ile Ile Gln Tyr Ala Pro Thr Ile Trp Val Arg Met Val Leu Met Thr  
2855 2860 2865

cac ttc ttc tcc att ctc atg gcc cag gac acc cta gac cag aac ctt 8995  
His Phe Phe Ser Ile Leu Met Ala Gln Asp Thr Leu Asp Gln Asn Leu  
2870 2875 2880 2885

aac ttt gaa atg tac gga tcg gtg tac tcc gtg agt cct ctg gac ctc 9043  
Asn Phe Glu Met Tyr Gly Ser Val Tyr Ser Val Ser Pro Leu Asp Leu  
2890 2895 2900

cca gcc ata att gaa agg tta cac ggg ctt gac gcc ttc tct ctg cac 9091

Pro Ala Ile Ile Glu Arg Leu His Gly Leu Asp Ala Phe Ser Leu His  
 2905 2910 2915

aca tac act ccc cac gaa ctg acg cgg gtg gct tca gcc ctc aga aaa 9139  
 Thr Tyr Thr Pro His Glu Leu Thr Arg Val Ala Ser Ala Leu Arg Lys  
 2920 2925 2930

ctt ggg gcg cca ccc ctc aga gcg tgg aag agt cgg gcg cgt gca gtt 9187  
 Leu Gly Ala Pro Pro Leu Arg Ala Trp Lys Ser Arg Ala Arg Ala Val  
 2935 2940 2945

agg gcg tcc ctc atc tcc cgt ggg ggg agg gcg gcc gtt tgc ggt cgg 9235  
 Arg Ala Ser Leu Ile Ser Arg Gly Gly Arg Ala Ala Val Cys Gly Arg  
 2950 2955 2960 2965

tac ctc ttc aac tgg gcg gtg aag acc aag ctc aaa ctc act cct ttg 9283  
 Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu Lys Leu Thr Pro Leu  
 2970 2975 2980

ccg gag gca cgc ctc ctg gat ttg tcc agt tgg ttt acc gtc ggc gcc 9331  
 Pro Glu Ala Arg Leu Leu Asp Leu Ser Ser Trp Phe Thr Val Gly Ala  
 2985 2990 2995

ggc ggg ggc gac att tat cac agc gtg tcg cgt gcc cga ccc cgc cta 9379  
 Gly Gly Gly Asp Ile Tyr His Ser Val Ser Arg Ala Arg Pro Arg Leu  
 3000 3005 3010

tta ctc ctt agc cta ctc cta ctt tct gta ggg gta ggc ctc ttc cta 9427  
 Leu Leu Leu Ser Leu Leu Leu Leu Ser Val Gly Val Gly Leu Phe Leu

3015

3020

3025

ctc ccc gct cga tag agcggcacac attagctaca ctccatagct aactgttcct 9482

Leu Pro Ala Arg

3030

tttttttttt tttttttttt tttttttttt tttttttctt tttttttttt tttccctctt 9542

tcttcccttc tcattctatt ctactttctt tcttggtggc tccatcttag ccctgggtcac 9602

ggctagctgt gaaaggtccg tgagccgcat gactgcagag agtgccgtaa ctggtctctc 9662

tgcagatcat gt

9674

&lt;210&gt; 6

&lt;211&gt; 3033

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 6

Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn

1 5 10 15

Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly

20 25 30

Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala

35 40 45

Thr Arg Lys Ala Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro

50 55 60

Ile Pro Lys His Arg Arg Ser Thr Gly Lys Ser Trp Gly Lys Pro Gly  
 65 70 75 80  
 Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp  
 85 90 95  
 Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Asn Asp Pro  
 100 105 110  
 Arg His Arg Ser Arg Asn Val Gly Lys Val Ile Asp Thr Leu Thr Cys  
 115 120 125  
 Gly Phe Ala Asp Leu Leu Gly Tyr Val Pro Val Val Gly Ala Pro Leu  
 130 135 140  
 Ser Gly Val Ala Ser Ala Leu Ala His Gly Val Arg Val Leu Glu Asp  
 145 150 155 160  
 Gly Val Asn Phe Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile  
 165 170 175  
 Phe Leu Leu Ala Leu Leu Ser Cys Ile Thr Thr Pro Val Ser Ala Val  
 180 185 190  
 Gln Val Lys Asn Thr Ser Asn Ala Tyr Met Ala Thr Asn Asp Cys Ser  
 195 200 205  
 Asn Asp Ser Ile Thr Trp Gln Leu Glu Ala Ala Val Leu His Val Pro  
 210 215 220  
 Gly Cys Val Pro Cys Glu Lys Met Gly Asn Thr Ser Arg Cys Trp Ile  
 225 230 235 240  
 Pro Val Ser Pro Asn Val Ala Val Arg Gln Pro Gly Ala Leu Thr Arg  
 245 250 255  
 Gly Leu Arg Thr His Ile Asp Met Val Val Leu Ser Ala Thr Leu Cys  
 260 265 270  
 Ser Ala Leu Tyr Val Gly Asp Leu Cys Gly Gly Val Met Leu Ala Ser  
 275 280 285  
 Gln Met Phe Ile Val Ser Pro Gln His His Trp Phe Val Gln Glu Cys

290 295 300  
Asn Cys Ser Ile Tyr Pro Gly Ala Ile Thr Gly His Arg Met Ala Trp  
305 310 315 320  
Asp Met Met Met Asn Trp Ser Pro Thr Thr Thr Met Ile Leu Ala Tyr  
325 330 335  
Val Met Arg Val Pro Glu Val Ile Ile Asp Ile Ile Ser Gly Ala His  
340 345 350  
Trp Gly Val Met Phe Gly Leu Ala Tyr Phe Ser Met Gln Gly Ala Trp  
355 360 365  
Ala Lys Val Val Val Ile Leu Leu Leu Ala Ser Gly Val Asp Ala Tyr  
370 375 380  
Thr Thr Thr Thr Gly Ser Ala Ala Gly Arg Thr Thr Ser Ser Leu Ala  
385 390 395 400  
Ser Ala Phe Ser Pro Gly Ala Arg Gln Asn Ile Gln Leu Ile Asn Thr  
405 410 415  
Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser  
420 425 430  
Leu His Thr Gly Phe Phe Thr Ala Leu Phe Tyr Ile His Lys Phe Asn  
435 440 445  
Ser Ser Gly Cys Pro Glu Arg Leu Ser Ala Cys Arg Asn Ile Glu Asp  
450 455 460  
Phe Arg Ile Gly Trp Gly Ala Leu Gln Tyr Asp Asp Asn Val Thr Asn  
465 470 475 480  
Pro Glu Asp Met Arg Pro Tyr Cys Trp His Tyr Pro Pro Lys Gln Cys  
485 490 495  
Gly Val Val Pro Ala Gly Thr Val Cys Gly Pro Val Tyr Cys Phe Thr  
500 505 510  
Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Leu Gly Val Pro Thr  
515 520 525

Tyr Thr Trp Gly Glu Asn Glu Thr Asp Val Phe Leu Leu Asn Ser Thr  
530 535 540

Arg Pro Pro Ser Gly Ser Trp Phe Gly Cys Thr Trp Met Asn Ser Thr  
545 550 555 560

Gly Phe Thr Lys Thr Cys Gly Ala Pro Pro Cys Arg Thr Arg Ala Asp  
565 570 575

Phe Asn Thr Ser Thr Asp Leu Leu Cys Pro Thr Asp Cys Phe Arg Lys  
580 585 590

His Pro Glu Ala Thr Tyr Ile Lys Cys Gly Ser Gly Pro Trp Leu Thr  
595 600 605

Pro Lys Cys Leu Val Asp Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys  
610 615 620

Thr Val Asn Tyr Ser Thr Phe Lys Ile Arg Met Tyr Val Gly Gly Val  
625 630 635 640

Glu His Arg Leu Met Ala Ala Cys Asn Phe Thr Arg Gly Asp Arg Cys  
645 650 655

Asn Leu Glu Asp Arg Asp Arg Ser Gln Gln Thr Pro Leu Leu His Ser  
660 665 670

Thr Thr Glu Trp Ala Ile Leu Pro Cys Ser Phe Ser Asp Leu Pro Ala  
675 680 685

Leu Ser Thr Gly Leu Leu His Leu His Gln Asn Ile Val Asp Val Gln  
690 695 700

Tyr Met Tyr Gly Leu Ser Pro Ala Leu Thr Gln Tyr Ile Val Arg Trp  
705 710 715 720

Glu Trp Val Val Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Val Cys  
725 730 735

Ala Cys Leu Trp Met Leu Ile Leu Leu Gly Gln Ala Glu Ala Ala Leu  
740 745 750

Glu Lys Leu Val Val Leu His Ala Ala Ser Ala Ala Ser Cys Asn Gly



755	760	765
Phe Leu Tyr Phe Val Ile Phe Leu Val Ala Ala Trp His Ile Lys Gly		
770	775	780
Arg Val Val Pro Leu Ala Ala Tyr Ser Leu Thr Gly Leu Trp Pro Phe		
785	790	795
Cys Leu Leu Leu Leu Ala Leu Pro Gln Gln Ala Tyr Ala Tyr Asp Ala		
805	810	815
Ser Val His Gly Gln Val Gly Ala Ala Leu Leu Val Leu Ile Thr Leu		
820	825	830
Phe Thr Leu Thr Pro Gly Tyr Lys Thr Leu Leu Ser Gln Ser Leu Trp		
835	840	845
Trp Leu Cys Tyr Leu Leu Thr Leu Ala Glu Thr Met Val Gln Glu Trp		
850	855	860
Ala Pro Ser Met Gln Ala Arg Gly Gly Arg Asp Gly Ile Ile Trp Ala		
865	870	875
Ala Thr Ile Phe Cys Pro Gly Val Val Phe Asp Ile Thr Lys Trp Leu		
885	890	895
Leu Ala Val Leu Gly Pro Gly Tyr Leu Leu Arg Gly Ala Leu Thr Arg		
900	905	910
Val Pro Tyr Phe Val Arg Ala His Ala Leu Leu Arg Met Cys Thr Met		
915	920	925
Val Arg His Leu Ala Gly Gly Arg Tyr Val Gln Met Ala Leu Leu Ala		
930	935	940
Leu Gly Arg Trp Thr Gly Thr Tyr Ile Tyr Asp His Leu Thr Pro Met		
945	950	955
Ser Asp Trp Ala Ala Ser Gly Leu Arg Asp Leu Ala Val Ala Val Glu		
965	970	975
Pro Ile Ile Phe Ser Pro Met Glu Lys Lys Val Ile Val Trp Gly Ala		
980	985	990

Glu Thr Ala Ala Cys Gly Asp Ile Leu His Gly Leu Pro Val Ser Ala  
995 1000 1005

Arg Leu Gly Arg Glu Ile Leu Leu Gly Pro Ala Asp Gly Tyr Thr Ser  
1010 1015 1020

Lys Gly Trp Lys Leu Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr  
1025 1030 1035 1040

Arg Gly Leu Leu Gly Ser Ile Val Val Ser Met Thr Gly Arg Asp Lys  
1045 1050 1055

Thr Glu Gln Ala Gly Glu Val Gln Val Leu Ser Thr Val Thr Gln Ser  
1060 1065 1070

Phe Leu Gly Thr Ser Ile Ser Gly Val Leu Trp Thr Val Tyr His Gly  
1075 1080 1085

Ala Gly Asn Lys Thr Leu Ala Gly Ser Arg Gly Pro Val Thr Gln Met  
1090 1095 1100

Tyr Ser Ser Ala Glu Gly Asp Leu Val Gly Trp Pro Ser Pro Pro Gly  
1105 1110 1115 1120

Thr Lys Ser Leu Glu Pro Cys Thr Cys Gly Ala Val Asp Leu Tyr Leu  
1125 1130 1135

Val Thr Arg Asn Ala Asp Val Ile Pro Ala Arg Arg Arg Gly Asp Lys  
1140 1145 1150

Arg Gly Ala Leu Leu Ser Pro Arg Pro Leu Ser Thr Leu Lys Gly Ser  
1155 1160 1165

Ser Gly Gly Pro Val Leu Cys Pro Arg Gly His Ala Val Gly Ile Phe  
1170 1175 1180

Arg Ala Ala Val Cys Ser Arg Gly Val Ala Lys Ser Ile Asp Phe Ile  
1185 1190 1195 1200

Pro Val Glu Thr Leu Asp Ile Val Thr Arg Ser Pro Thr Phe Ser Asp  
1205 1210 1215

Asn Ser Thr Pro Pro Ala Val Pro Gln Thr Tyr Gln Val Gly Tyr Leu

1220 1225 1230  
His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro Val Ala Tyr  
1235 1240 1245  
Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala  
1250 1255 1260  
Thr Leu Gly Phe Gly Ala Tyr Leu Ser Lys Ala His Gly Ile Asn Pro  
1265 1270 1275 1280  
Asn Ile Arg Thr Gly Val Arg Thr Val Thr Thr Gly Glu Pro Ile Thr  
1285 1290 1295  
Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys Ala Gly Gly  
1300 1305 1310  
Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser Val Asp Ala Thr  
1315 1320 1325  
Thr Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly  
1330 1335 1340  
Val Arg Leu Thr Val Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr  
1345 1350 1355 1360  
Thr Pro His Pro Asn Ile Glu Glu Val Ala Leu Gly Gln Glu Gly Glu  
1365 1370 1375  
Ile Pro Phe Tyr Gly Arg Ala Phe Pro Leu Ser Tyr Ile Lys Gly Gly  
1380 1385 1390  
Arg His Leu Ile Phe Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala  
1395 1400 1405  
Thr Ala Leu Arg Gly Met Gly Leu Asn Ala Val Ala Tyr Tyr Arg Gly  
1410 1415 1420  
Leu Asp Val Ser Ile Ile Pro Thr Gln Gly Asp Val Val Val Val Ala  
1425 1430 1435 1440  
Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile  
1445 1450 1455

Asp Cys Asn Val Ala Val Thr Gln Ala Val Asp Phe Ser Leu Asp Pro  
1460 1465 1470  
Thr Phe Thr Ile Thr Thr Gln Thr Val Pro Gln Asp Ala Val Ser Arg  
1475 1480 1485  
Ser Gln Arg Arg Gly Arg Thr Gly Arg Gly Arg Leu Gly Ile Tyr Arg  
1490 1495 1500  
Tyr Val Ser Thr Gly Glu Arg Ala Ser Gly Met Phe Asp Ser Val Val  
1505 1510 1515 1520  
Leu Cys Glu Cys Tyr Asp Ala Gly Ala Ala Trp Tyr Glu Leu Ser Pro  
1525 1530 1535  
Val Glu Thr Thr Val Arg Leu Arg Ala Tyr Phe Asn Thr Pro Gly Leu  
1540 1545 1550  
Pro Val Cys Gln Asp His Leu Glu Phe Trp Glu Ala Val Phe Thr Gly  
1555 1560 1565  
Leu Thr His Ile Asp Ala His Phe Leu Ser Gln Thr Lys Gln Ser Gly  
1570 1575 1580  
Glu Asn Phe Ala Tyr Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg  
1585 1590 1595 1600  
Ala Lys Ala Pro Pro Pro Ser Trp Asp Val Met Trp Lys Cys Leu Thr  
1605 1610 1615  
Arg Leu Lys Pro Thr Leu Val Gly Pro Thr Pro Leu Leu Tyr Arg Leu  
1620 1625 1630  
Gly Ser Val Thr Asn Glu Val Thr Leu Thr His Pro Val Thr Lys Tyr  
1635 1640 1645  
Ile Ala Thr Cys Met Gln Ala Asp Leu Glu Val Met Thr Ser Thr Trp  
1650 1655 1660  
Val Leu Ala Gly Gly Val Leu Ala Ala Val Ala Ala Tyr Cys Leu Ala  
1665 1670 1675 1680  
Thr Gly Cys Val Ser Ile Ile Gly Arg Leu His Ile Asn Gln Arg Ala

1685 1690 1695  
Val Val Ala Pro Asp Lys Glu Val Leu Tyr Glu Ala Phe Asp Glu Met  
1700 1705 1710  
Glu Glu Cys Ala Ser Arg Ala Ala Leu Leu Glu Glu Gly Gln Arg Ile  
1715 1720 1725  
Ala Glu Met Leu Lys Ser Lys Ile Gln Gly Leu Leu Gln Gln Ala Ser  
1730 1735 1740  
Lys Gln Ala Gln Asp Ile Gln Pro Ala Val Gln Ala Ser Trp Pro Lys  
1745 1750 1755 1760  
Met Glu Gln Phe Trp Ala Lys His Met Trp Asn Phe Ile Ser Gly Ile  
1765 1770 1775  
Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn Pro Ala Val Ala  
1780 1785 1790  
Ser Met Met Ala Phe Ser Ala Ala Leu Thr Ser Pro Leu Ser Thr Ser  
1795 1800 1805  
Thr Thr Ile Leu Leu Asn Ile Leu Gly Gly Trp Leu Ala Ser Gln Ile  
1810 1815 1820  
Ala Pro Pro Ala Gly Ala Thr Gly Phe Val Val Ser Gly Leu Val Gly  
1825 1830 1835 1840  
Ala Ala Val Gly Ser Ile Gly Leu Gly Lys Val Leu Val Asp Ile Leu  
1845 1850 1855  
Ala Gly Tyr Gly Ala Gly Ile Ser Gly Ala Leu Val Ala Phe Lys Ile  
1860 1865 1870  
Met Ser Gly Glu Lys Pro Ser Met Glu Asp Val Ile Asn Leu Leu Pro  
1875 1880 1885  
Gly Ile Leu Ser Pro Gly Ala Leu Val Val Gly Val Ile Cys Ala Ala  
1890 1895 1900  
Ile Leu Arg Arg His Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met  
1905 1910 1915 1920

Asn Arg Leu Ile Ala Phe Ala Ser Arg Gly Asn His Val Ala Pro Thr  
1925 1930 1935  
His Tyr Val Thr Glu Ser Asp Ala Ser Gln Arg Val Thr Gln Leu Leu  
1940 1945 1950  
Gly Ser Leu Thr Ile Thr Ser Leu Leu Arg Arg Leu His Asn Trp Ile  
1955 1960 1965  
Thr Glu Asp Cys Pro Ile Pro Cys Ala Gly Ser Trp Leu Arg Asp Val  
1970 1975 1980  
Trp Asp Trp Val Cys Thr Ile Leu Thr Asp Phe Lys Asn Trp Leu Thr  
1985 1990 1995 2000  
Ser Lys Leu Phe Pro Lys Met Pro Gly Leu Pro Phe Ile Ser Cys Gln  
2005 2010 2015  
Lys Gly Tyr Lys Gly Val Trp Ala Gly Thr Gly Ile Met Thr Thr Arg  
2020 2025 2030  
Cys Pro Cys Gly Ala Asn Ile Ser Gly Asn Val Arg Leu Gly Ser Met  
2035 2040 2045  
Arg Ile Thr Gly Pro Lys Thr Cys Met Asn Thr Trp Gln Gly Thr Phe  
2050 2055 2060  
Pro Ile Asn Cys Tyr Thr Glu Gly Gln Cys Leu Pro Lys Pro Ala Leu  
2065 2070 2075 2080  
Asn Phe Lys Thr Ala Ile Trp Arg Val Ala Ala Ser Glu Tyr Ala Glu  
2085 2090 2095  
Val Thr Gln His Gly Ser Tyr Ala Tyr Ile Thr Gly Leu Thr Thr Asp  
2100 2105 2110  
Asn Leu Lys Val Pro Cys Gln Leu Pro Ser Pro Glu Phe Phe Ser Trp  
2115 2120 2125  
Val Asp Gly Val Gln Ile His Arg Ser Ala Pro Thr Pro Lys Pro Phe  
2130 2135 2140  
Phe Arg Asp Glu Val Ser Phe Ser Val Gly Leu Asn Ser Phe Val Val

2145                      2150                      2155                      2160  
Gly Ser Gln Leu Pro Cys Asp Pro Glu Pro Asp Thr Glu Val Val Met  
                         2165                      2170                      2175  
Ser Met Leu Thr Asp Pro Ser His Ile Thr Ala Glu Ala Ala Ala Arg  
                         2180                      2185                      2190  
Arg Leu Ala Arg Gly Ser Pro Pro Ser Glu Ala Ser Ser Ser Ala Ser  
                         2195                      2200                      2205  
Gln Leu Ser Ala Pro Ser Leu Arg Ala Thr Cys Thr Thr His Gly Arg  
                         2210                      2215                      2220  
Thr Tyr Asp Val Asp Met Val Asp Ala Asn Leu Phe Met Gly Gly Gly  
2225                      2230                      2235                      2240  
Val Ile Arg Ile Glu Ser Glu Ser Lys Val Val Val Leu Asp Ser Leu  
                         2245                      2250                      2255  
Asp Ser Met Thr Glu Glu Glu Gly Asp Leu Glu Pro Ser Val Pro Ser  
                         2260                      2265                      2270  
Glu Tyr Met Leu Pro Arg Lys Arg Phe Pro Pro Ala Leu Pro Ala Trp  
                         2275                      2280                      2285  
Ala Arg Pro Asp Tyr Asn Pro Pro Leu Val Glu Ser Trp Lys Arg Pro  
                         2290                      2295                      2300  
Asp Tyr Gln Pro Pro Thr Val Ala Gly Cys Ala Leu Pro Pro Pro Lys  
2305                      2310                      2315                      2320  
Lys Thr Pro Thr Pro Pro Pro Arg Arg Arg Arg Thr Val Gly Leu Ser  
                         2325                      2330                      2335  
Glu Ser Thr Ile Gly Asp Ala Leu Gln Gln Leu Ala Ile Lys Ser Phe  
                         2340                      2345                      2350  
Gly Gln Pro Pro Pro Ser Gly Asp Ser Gly Leu Ser Thr Gly Ala Asp  
                         2355                      2360                      2365  
Ala Ala Asp Ser Gly Asp Arg Thr Pro Pro Asp Glu Leu Ala Leu Ser  
                         2370                      2375                      2380

Glu Thr Gly Ser Thr Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly  
2385 2390 2395 2400  
Asp Pro Asp Leu Glu Pro Glu Gln Val Glu Leu Gln Pro Pro Pro Gln  
2405 2410 2415  
Gly Gly Glu Ala Ala Pro Gly Ser Asp Ser Gly Ser Trp Ser Thr Cys  
2420 2425 2430  
Ser Glu Glu Asp Asp Ser Val Val Cys Cys Ser Met Ser Tyr Ser Trp  
2435 2440 2445  
Thr Gly Ala Leu Ile Thr Pro Cys Ser Pro Glu Glu Glu Lys Leu Pro  
2450 2455 2460  
Ile Asn Ser Leu Ser Asn Ser Leu Leu Arg Tyr His Asn Lys Val Tyr  
2465 2470 2475 2480  
Cys Thr Thr Ser Lys Ser Ala Ser Leu Arg Ala Lys Lys Val Thr Phe  
2485 2490 2495  
Asp Arg Met Gln Val Leu Asp Ala Tyr Tyr Asp Ser Val Leu Lys Asp  
2500 2505 2510  
Ile Lys Leu Ala Ala Ser Lys Val Ser Ala Arg Leu Leu Thr Leu Glu  
2515 2520 2525  
Glu Ala Cys Gln Leu Thr Pro Pro His Ser Ala Arg Ser Lys Tyr Gly  
2530 2535 2540  
Phe Gly Ala Lys Glu Val Arg Ser Leu Ser Gly Arg Ala Val Asn His  
2545 2550 2555 2560  
Ile Lys Ser Val Trp Lys Asp Leu Leu Glu Asp Ser Gln Thr Pro Ile  
2565 2570 2575  
Pro Thr Thr Ile Met Ala Lys Asn Glu Val Phe Cys Val Asp Pro Ala  
2580 2585 2590  
Lys Gly Gly Lys Lys Pro Ala Arg Leu Ile Val Tyr Pro Asp Leu Gly  
2595 2600 2605  
Val Arg Val Cys Glu Lys Met Ala Leu Tyr Asp Val Thr Gln Lys Leu



2610                      2615                      2620  
Pro Gln Ala Val Met Gly Ala Ser Tyr Gly Phe Gln Tyr Ser Pro Ala  
2625                      2630                      2635                      2640  
Gln Arg Val Glu Phe Leu Leu Lys Ala Trp Ala Glu Lys Arg Asp Pro  
                         2645                      2650                      2655  
Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe Asp Ser Thr Val Thr Glu  
                         2660                      2665                      2670  
Arg Asp Ile Arg Thr Glu Glu Ser Ile Tyr Gln Ala Cys Ser Leu Pro  
                         2675                      2680                      2685  
Glu Glu Ala Arg Thr Ala Ile His Ser Leu Thr Glu Arg Leu Tyr Val  
                         2690                      2695                      2700  
Gly Gly Pro Met Phe Asn Ser Lys Gly Gln Ser Cys Gly Tyr Arg Arg  
2705                      2710                      2715                      2720  
Cys Arg Ala Ser Gly Val Leu Thr Thr Ser Met Gly Asn Thr Ile Thr  
                         2725                      2730                      2735  
Cys Tyr Val Lys Ala Leu Ala Ala Cys Lys Ala Ala Gly Ile Ile Ala  
                         2740                      2745                      2750  
Pro Thr Met Leu Val Cys Gly Asp Asp Leu Val Val Ile Ser Glu Ser  
                         2755                      2760                      2765  
Gln Gly Thr Glu Glu Asp Glu Arg Asn Leu Arg Ala Phe Thr Glu Ala  
                         2770                      2775                      2780  
Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp Pro Pro Arg Pro Glu Tyr  
2785                      2790                      2795                      2800  
Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser Asn Val Ser Val Ala Leu  
                         2805                      2810                      2815  
Gly Pro Gln Gly Arg Arg Arg Tyr Tyr Leu Thr Arg Asp Pro Thr Thr  
                         2820                      2825                      2830  
Ser Ile Ala Arg Ala Ala Trp Glu Thr Val Arg His Ser Pro Val Asn  
                         2835                      2840                      2845

Ser Trp Leu Gly Asn Ile Ile Gln Tyr Ala Pro Thr Ile Trp Val Arg  
2850 2855 2860  
Met Val Leu Met Thr His Phe Phe Ser Ile Leu Met Ala Gln Asp Thr  
2865 2870 2875 2880  
Leu Asp Gln Asn Leu Asn Phe Glu Met Tyr Gly Ser Val Tyr Ser Val  
2885 2890 2895  
Ser Pro Leu Asp Leu Pro Ala Ile Ile Glu Arg Leu His Gly Leu Asp  
2900 2905 2910  
Ala Phe Ser Leu His Thr Tyr Thr Pro His Glu Leu Thr Arg Val Ala  
2915 2920 2925  
Ser Ala Leu Arg Lys Leu Gly Ala Pro Pro Leu Arg Ala Trp Lys Ser  
2930 2935 2940  
Arg Ala Arg Ala Val Arg Ala Ser Leu Ile Ser Arg Gly Gly Arg Ala  
2945 2950 2955 2960  
Ala Val Cys Gly Arg Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu  
2965 2970 2975  
Lys Leu Thr Pro Leu Pro Glu Ala Arg Leu Leu Asp Leu Ser Ser Trp  
2980 2985 2990  
Phe Thr Val Gly Ala Gly Gly Gly Asp Ile Tyr His Ser Val Ser Arg  
2995 3000 3005  
Ala Arg Pro Arg Leu Leu Leu Leu Ser Leu Leu Leu Leu Ser Val Gly  
3010 3015 3020  
Val Gly Leu Phe Leu Leu Pro Ala Arg  
3025 3030

&lt;210&gt; 7

&lt;211&gt; 8024

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: replicon

<400> 7

accugccccc aaauaggggcg acacuccgcc augaaucacu cccugugag gaacuacugu 60  
cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120  
ccccuucccg ggagagccau aguggucugc ggaaccggug aguacaccgg aaaugccggg 180  
aagacugggu ccuuucuugg auaaaccac ucuaugcccg gccauuuggg cgugcccccg 240  
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300  
cgcuugcgag ugccccggga ggucucguag accgugcacc augagcaca auccuaaacc 360  
ucaaagaaaa accaaaagaa acaccaaccg ugcaccaaug auugaacaag auggauugca 420  
cgcagguucu ccggccgcuu ggguggagag gcuauucggc uaugacuggg cacaacagac 480  
aaucggcugc ucugaugccg ccguguuccg gcugucagcg caggggcgcc cgguucuuuu 540  
ugucaagacc gaccuguccg gugccugaa ugaacugcag gacgaggcag cgcggcuauc 600  
guggcuggcc acgacgggcg uuccuugcgc agcugugcuc gacguuguca cugaagcggg 660  
aagggaugcug cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720  
uccugccgag aaagaucca ucauggcuga ugcaaugcgg cggcugcaua cgcuugaucc 780  
ggcuaccugc ccuucgacc accaagcgaa acaucgauc gagcgagcac guacucggau 840  
ggaagccggu cuugucgauc aggaugaucu ggacgaagag caucaggggc ucgcgccagc 900  
cgaacuguuc gccaggcuca aggcgcgcgau gcccgcggc gaggaucucg ucgugaccca 960  
uggcgaugcc ugcuuuccga auaucauggu ggaaauggc cgcuuuucug gauucaucga 1020  
cuguggccgg cugggugugg cggaccgcu ucaggacaua gcguuggcua cccgugauau 1080  
ugcugaagag cuuggcggcg aaugggcuga ccguuccuc gugcuuuacg guaucgccgc 1140  
ucccgauucg cagcgcaucg ccuucuaucg ccuucugac gaguucuucu gaguuuaaac 1200  
ccucuccuc ccccccccu aacguuacug gccgaagccg cuuggaauaa ggccggugug 1260  
cguuugucua uauguuuuu uccaccauau ugccgucuuu uggcaaugug agggcccggg 1320

aaccuggccc ugucuucuug acgagcauuc cuaggggucu uuccccucuc gccaaaggaa 1380  
ugcaaggucu guugaauguc gugaaggaag caguuccucu ggaagcuucu ugaagacaaa 1440  
caacgucugui agcgacccuu ugcaggcagc ggaaccccc accuggcgac aggugccucu 1500  
gcggccaaaa gccacgugua uaagauacac cugcaaaggc ggcacaaccc cagugccacg 1560  
uugugaguug gauaguugug gaaagaguca aauggcucuc cucaagcgua uucaacaagg 1620  
ggcugaagga ugcccagaag guaccccauu guaugggauc ugaucugggg ccucggugca 1680  
caugcuuuac auguguuuag ucgagguuaa aaaaacgucu agggcccccg aaccacgggg 1740  
acgugguuuu ccuuugaaaa acacgaugau accauggcuc ccaucacugc uuaugcccag 1800  
caaacacgag gccuccuggg cgccauagug gugaguauga cggggcguga caggacagaa 1860  
caggccgggg aaguccaaau ccuguccaca gucucucagu ccuuccucgg aacaaccauc 1920  
ucggggguuu uguggacugu uuaccacgga gcuggcaaca agacucuagc cggcuuacgg 1980  
gguccgguca cgcagaugua cucgagugcu gagggggacu ugguaggcug gccagcccc 2040  
ccugggacca agucuugga gccgugcaag uguggagccg ucgaccuaua ucuggucacg 2100  
cggaacgcug augucauccc ggcucggaga cgcggggaca agcggggagc auugcucucc 2160  
ccgagacca uuucgaccuu gaaggggucc ucgggggggc cggugcucug ccuaggggc 2220  
cacgucguug ggcucuuccg agcagcugug ugcucucggg gcguggccaa auccaucgau 2280  
uucauccccg uugagacacu cgacguuguu acaaggucuc ccacuuucag ugacaacagc 2340  
acgccaccgg cugugcccca gaccuauacg gucggguacu ugcaugcucc aacuggcagu 2400  
ggaaagagca ccaagguccc ugucgcguau gccgccagg gguacaaagu acuagugcuu 2460  
aaccucccgg uagcugccac ccugggguuu ggggcguacc uauccaaggc acauggcauc 2520  
aaucccaaca uuaggacugg agucaggacc gugaugaccg gggaggccau cacguacucc 2580  
acauauggca aaauucucgc cgaugggggc ugcgcuaagc gcgccuauga caucaua 2640  
ugcgaugaau gccacgcugu ggaugcuacc uccauucucg gcaucggaac gguccuugau 2700  
caagcagaga cagccggggu cagacuaacu gucguggcua cggccacacc ccccggguca 2760  
gugacaaccc cccauccga uauagaaaag guaggccucg ggcgggaggg ugagaucucc 2820  
uucuauggga gggcgauucc ccuauccugc aucaaggag ggagacaccu gauuuucugc 2880  
cacucaaaga aaaaguguga cgagcucgcg gcggcccuuc ggggcauggg cuugaaugcc 2940  
guggcauacu auagaggguu ggacgucucc auaauaccag cucagggaga uguggugguc 3000  
gucgccaccg acgccucau gacgggguac acuggagacu uugacuccgu gaucgacugc 3060

aauguagcgg ucaccaagc ugucgacuuc agccuggacc ccaccuucac uauaaccaca 3120  
cagacugucc cacaagacgc ugucucacgc agucagcgcc gcggg'gcac agguagagga 3180  
agacagggca cuuauaggua uguuuccacu ggugaacgag ccucaggaau guuugacagu 3240  
guagugcuuu gugagugcua cgacgcaggg gcugcguggu acgaucucac accagcggag 3300  
accaccguca ggcuuagagc guauuucaac acgcccggcc uacccgugug ucaagaccau 3360  
cuugaauuuu gggaggcagu uuucaccggc cucacacaca uagacgcca cuuccucucc 3420  
caaacaaagc aagcggggga gaacuucgcg uaccuaguag ccuaccaagc uacggugugc 3480  
gccagagcca agggccucc cccguccugg gacgccaugu ggaagugccu ggcccgcacuc 3540  
aagccuacgc uugcggggcc cacaccucuc cuguaccguu uggggccuau uaccaaugag 3600  
gucaccuca cacaccugg gacgaaguac aucgccacau gcaugcaagc ugaccuugag 3660  
gucaugacca gcacgugggu ccuagcugga ggaguccugg cagccgucgc cgcauuugc 3720  
cuggcgacug gaugcguuuc caucaucggc cgcuugcacg ucaaccagcg agucgucguu 3780  
gcgcccgaaua aggagguccu gnaugaggcu uuugaugaga uggaggaaug cgccucuagg 3840  
gcggcucuca ucgaagaggg gcagcggaua gccgagaugu ugaaguccaa gauccaaggc 3900  
uugcugcagc agggccucuaa gcaggcccag gacauacaac ccgcuaugca ggcuucaugg 3960  
cccaaagugg aacaauuuug ggccagacac auguggaacu ucauuagcgg cauccauuac 4020  
cucgcaggau ugucaacacu gccagggaac cccgcggugg cuuccaugau ggcauucagu 4080  
gccgcccuca ccaguccguu gucgaccagu accaccaucc uucucaacau caugggaggc 4140  
ugguuagcgu cccagaucgc accaccgcg gggggccaccg gcuuugucgu caguggccug 4200  
gugggggucg ccgugggcag cauaggccug gguaaggugc ugguggacau ccuggcagga 4260  
uauggugcgg gcauuucggg ggcccucguc gcauucaaga ucaugucugg cgagaagccc 4320  
ucuauggaag augucauaa ucuacugccu gggauccugu cuccgggagc ccugguggug 4380  
ggggucaucu gcgcggccau ucugcgccgc cacgugggac cgggggaggg gcgguccaa 4440  
uggaugaaca ggcuuauugc cuuugcuucc agaggaaacc acgucgcccc uacucacuac 4500  
gugacggagu cggaugcgc gcagcgugug acccaacuac uuggcucucu uacuauaacc 4560  
agccuacuca gaagacucca caauuggaua acugaggacu gcccacucc augcuccgga 4620  
uccuggcucc gcgacgugug ggacuggguu ugcaccaucu ugacagacuu caaaaauugg 4680  
cugaccucua aauguuucc caagcugccc ggccucccu ucaucucuug ucaaaagggg 4740  
uacaagggug ugugggcccg cacuggcauc augaccacgc gcugcccuug cggcgccaac 4800

aucucuggca auguccgccu gggcucuau aggaucacag ggccuaaaac cugcaugaac 4860  
accuggcagg ggaccuuucc uaucaauugc uacacggagg gccagugcgc gccgaaaccc 4920  
cccacgaacu acaagaccgc caucuggagg guggcggccu cggaguacgc ggaggugacg 4980  
cagcaugggu cguacuccua uguaacagga cugaccacug acaaucugaa aaauccuugc 5040  
caacuaccuu cuccagaguu uuucuccugg guggacggug ugcagaucca uagguuugca 5100  
cccacaccaa agccguuuuu ccgggaugag gucucguucu gcguugggcu uaauuccuau 5160  
gcugucgggu cccagcuucc cugugaaccu gagcccgacg cagacguauu gagguccaug 5220  
cuaacagauc cgccccacau cacggcggag acugcggcgc ggcgcuuggc acggggauca 5280  
ccuccaucug aggcgagcuc cucagugagc cagcuauacg caccgucgcu gcggggccacc 5340  
ugcaccaccc acagcaacac cuaugacgug gacauggucg augccaaccu gcucauggag 5400  
ggcggugugg cucagacaga gccugagucc agggugcccg uucuggacuu ucucgagcca 5460  
auggccgagg aagagagcga ccuugagccc ucaauaccuau cggagugcau gcuccccagg 5520  
agcggguuuc cacgggccuu accggcuugg gcacggccug acuacaacc gccgcucgug 5580  
gaaucgugga ggaggccaga uuaccaaccg cccaccguug cugguugugc ucucgggggg 5640  
cccaagaagg ccccgacgcc ucccccaagg agacgccgga cagugggucu gagcgagagc 5700  
accauaucag aagcccuca gcaacuggcc aucaagaccu uggccagcc cccucgagc 5760  
ggugaugcag gcucguccac gggggcgggc gccgccgaau ccggcggucc gacgucccu 5820  
ggugagccgg cccucacaga gacagguucc gccuccucua ugccccccu cgaggggggag 5880  
ccuggagauc cggaccugga gucugaucag guagagcuuc aaccucggc ccaggggggg 5940  
gggguagcuc ccgguucggg cucggggucu uggucuacuu gcuccgagga ggacgauacc 6000  
accgugugcu gcuccauguc auacuccugg accggggcuc uaauaacucc cuguagcccc 6060  
gaagaggaaa aguugccaau caaccuuug aguaacucgc uguugcgaua ccauaacaag 6120  
guguacugua caacaucaaa gagcgccuca cagagggcua aaaagguaac uuugacagg 6180  
acgcaagugc ucgacgcca uuaugacuca gucuuaaagg acaucaagcu agcggcuucc 6240  
aaggucagcg caaggcuccu caccuuggag gaggcgugcc aguugacucc accccauucu 6300  
gcaagaucca aguauggauu cggggccaag gagguccgca gcuuguccgg gagggccguu 6360  
aaccacauca aguccgugug gaaggaccuc cuggaagacc cacaacacc aaaucccaca 6420  
accaucaugg ccaaaauga gguguucugc guggaccccg ccaagggggg uaagaaacca 6480  
gcucgccuca ucguuuacc ugaccucggc guccgggucu gcgagaaaau ggcccucua 6540

gacauuacac aaaagcuucc ucaggcggua augggagcuu ccuauggcuu ccaguacucc 6600  
 ccugcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagga cccaugggu 6660  
 uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720  
 gaguccauau accaggccug cucccugccc gagggaggccc gcacugccau acacucgcug 6780  
 acugagagac uuuacguagg agggcccaug uucaacagca agggucuaac cugcgguuac 6840  
 agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccu cacaugcuau 6900  
 gugaaagccc uagcggccug caaggcugcg gggauaguug cgcccacaau gcugguauugc 6960  
 ggcaaugacc uaguagucan cucagaaagc caggggacug agggaggacga gcggaaccug 7020  
 agagccuua cggaggccau gaccagguac ucugccccuc cuggugaucc ccccagaccg 7080  
 gaauaugacc uggagcuau aacauccugu uccucaaau ugucuguggc guugggccc 7140  
 cggggccgcc gcagauacua ccugaccaga gacccaacca cuccacucgc ccgggcugcc 7200  
 ugggaaacag uuagacacuc ccuaucaau ucauggcugg gaaacaucau ccaguaugcu 7260  
 ccaaccuau gguuucgcau gguccuaau acacacuuc ucuccauuc caugguccaa 7320  
 gacaccucgg accagaaccu caacuugag auguauugau caguauacuc cgugaauccu 7380  
 uuggaccuuc cagccauau ugagagguua cacgggcuug acgccuuuc uaugcacaca 7440  
 uacucucacc acgaacugac gcggguggcu ucagcccua gaaaacuugg ggcgccacc 7500  
 cucagggugu ggaagagucg ggcucgcgc gucagggcg cccucaucuc ccuggaggg 7560  
 aaagcggccg uuugcggccg auaucucuuc aaugggcgg ugaagaccaa gcucuaacuc 7620  
 acuccauugc cggaggcgcg ccuacuggac uuauccagu gguuaccgu cggcgccggc 7680  
 gggggcgaca uuuuucacag cgugucgcgc gcccgaaccc gcucauuacu cuucggccua 7740  
 cuccuacuu ucguaggggu aggccucuuc cuacuccccg cucgguagag cggcacacac 7800  
 uagguacacu ccuagcuua cuguuccuu uuuuuuuuu uuuuuuuuu uuuuuuuuu 7860  
 uuuuuuuuu cuuuuuuuuu uuuuuccuc uuucuuucc ucucauuua uucuaucuu 7920  
 uuucuuugg gcuccauuu agcccuaguc acggcuagcu gugaaagguc cgugagccgc 7980  
 augacugcag agagugccgu aacuggucuc ucugcagau augu 8024

&lt;210&gt; 8

&lt;211&gt; 7994

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: replicon

<400> 8

accugccccc aaauaggggcg acacuccgcc augaaucacu cccugugag gaacuacugu 60  
cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120  
ccccucuccg ggagagccau aguggucugc ggaaccggug aguacaccgg aaaugccggg 180  
aagacugggu ccuucuuugg auaaaccac ucuaugccg gccauuuggg cgugcccccg 240  
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300  
cgcuugcgag ugccccggga ggucucguag accgugcacc augagcaca auccuaaacc 360  
ucaaagaaaa accaaaagaa acaccaaccg ugcaccaaug auugaacaag auggauugca 420  
cgcagguucu ccggccgcuu ggguggagag gcuaauccggc uaugacuggg cacaacagac 480  
aaucggcugc ucugaugccg ccguguuccg gcugucagcg caggggcgcc cgguuuuuu 540  
ugucaagacc gaccuguccg gugcccugaa ugaacugcag gacgaggcag cgcggcuau 600  
guggcuggcc acgacgggcg uuccuugcgc agcugugcuc gacguuguca cugaagcggg 660  
aagggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720  
uccugccgag aaaguaucca ucauggcuga ugcaaugcgg cggcugcaua cguugaucc 780  
ggcuaccugc ccuucgacc accaagcgaa acaucgauc gagcgagcac guacucggau 840  
ggaagccggu cuugucgauc aggaugaucu ggacgaagag caucaggggc ucgcgccagc 900  
cgaacuguuc gccaggcuca aggcgcgcau gcccgacggc gaggaucucg ucgugacca 960  
uggcgaugcc ugcuugccga auaucauggu ggaaaauaggc cgcuuuucug gauucaucga 1020  
cuguggccgg cugggugugg cggaccgcu ucaggacaua gcguuggcu cccgugauau 1080  
ugcugaagag cuuggcggcg aaugggcuga ccgcuuccuc gugcuuuacg guaucgccgc 1140  
ucccgauucg cagcgcaucg ccuucuaucg ccuucugac gaguucuucu gaguuuaaac 1200  
ccucuccuc ccccccccu aacguuacug gccgaagccg cuuggaauaa ggccggugug 1260  
cguuugucua uauguuuuu uccaccauau ugccgucuuu uggcaaugug agggcccgga 1320



aaccuggccc ugucuucuug acgagcauuc cuaggggucu uuccccucuc gccaaaggaa 1380  
ugcaaggucu guugaauguc gugaaggaag caguuccucu ggaagcuucu ugaagacaaa 1440  
caacgucugu agcgacccuu ugcaggcagc ggaaccccc accuggcgac aggugccucu 1500  
gcggccaaaa gccacgugua uaagauacac cugcaaaggc ggcacaaccc cagugccacg 1560  
uugugaguug gauaguugug gaaagaguca aauggcucuc cucaagcgua uucaacaagg 1620  
ggcugaagga ugcccagaag guaccccauu guaugggauc ugaucugggg ccucggugca 1680  
caugcuuuac auguguuuag ucgagguuaa aaaaacgucu agggcccccg aaccacgggg 1740  
acuggguuuu ccuuugaaaa acacgaugau accauggcuc ccaucacugc uuaucccag 1800  
caaacacgag gccuccuggg cgccauagug gugaguauga cggggcguga caggacagaa 1860  
caggccgggg aaguccaaau ccuguccaca gucucucagu ccuuccucgg aacaaccauc 1920  
ucggggguuu uguggacugu uuaccacgga gcuggcaaca agacucuagc cggcuuacgg 1980  
gguccggua cgcagaugua cucgagugcu gagggggacu ugguaggcug gccagcccc 2040  
ccugggacca agucuugga gccgugcaag uguggagccg ucgaccuaua ucuggucacg 2100  
cggaacgcug augucauccc ggcucggaga cgcggggaca agcggggagc auugcucucc 2160  
ccgagacca uuucgaccuu gaaggggucc ucgggggggc cggugcucug ccuaggggc 2220  
cacgucguug ggcucuuccg agcagcugug ugcucucggg gcguggccaa auccaucgau 2280  
uucaucccg uugagacacu cgacguuguu acaaggucuc ccacuuucag ugacaacagc 2340  
acgccaccgg cugugccca gaccuauacg gucggguacu ugcaugcucc aacuggcagu 2400  
ggaaagagca ccaagguccc ugucgcguau gccgccagg gguacaaagu acuagugcuu 2460  
aaccucugg uagcugccac ccugggguuu ggggcguacc uauccaaggc acauggcauc 2520  
aaucccaaca uuaggacugg agucaggacc gugaugaccg gggaggccau cacguacucc 2580  
acauauggca aaauucucgc cgaugggggc ugcguagcg gcgccuaga caucaua 2640  
ugcgaugaau gccacgcugu ggaugcuacc uccauucucg gcaucggaac gguccuugau 2700  
caagcagaga cagccggggu cagacuaacu gucuggcua cgccacacc cccggguca 2760  
gugacaaccc ccaucccg uauagaagag guaggccucg ggcgggaggg ugagaucucc 2820  
uucuauggga gggcgauucc ccuauccugc aucaaggag ggagacaccu gauuuucugc 2880  
cacucaaaga aaaaguguga cgagcucgcg gcggcccuuc ggggcauggg cuugaaugcc 2940  
guggcauacu auagaggguu ggacgucucc auauaccag cucagggaga ugugguguc 3000  
gucgccaccg acggccuau gacgggguac acuggagacu uugacuccgu gaucgacugc 3060

aauguagcgg ucaccaagc ugucgacuuc agccuggacc ccaccuucac uauaaccaca 3120  
cagacugucc cacaagacgc ugucucacgc agucagcgcc gcggg'gcac agguagagga 3180  
agacagggca cuuauaggua uguuuccacu ggugaacgag ccucaggaau guuugacagu 3240  
guagugcuuu gugagugcua cgacgcaggg gcugcguggu acgaucucac accagcggag 3300  
accaccguca ggcuuagagc guauuucaac acgcccggcc uacccgugug ucaagaccau 3360  
cuugaauuuu gggaggcagu uuucaccggc cucacacaca uagacgcca cuuccucucc 3420  
caaacaaagc aagcggggga gaacuucgcg uaccuaguag ccuaccaagc uacggugugc 3480  
gccagagcca aggccccucc cccguccugg gacgccaugu ggaagugccu ggcccgcacuc 3540  
aagccuacgc uugcggggccc cacaccucuc cuguaccguu ugaggccuau uaccaaugag 3600  
gucacccuca cacaccugg gacgaaguac aucgccacau gcaugcaagc ugaccuugag 3660  
gucaugacca gcacgugggu ccuagcugga ggaguccugg cagccgucgc cgcauuugc 3720  
cuggcgacug gaugcguuuc caucaucggc cgcuuacag ucaaccagcg agucgucguu 3780  
gcgcccgaaua aggagguccu guaugaggcu uuugaugaga ugagggaug cgccucuagg 3840  
gcggcucuca ucgaagaggg gcagcggaaua gccgagaugu ugaaguccaa gauccaaggc 3900  
uugcugcagc aggccucuaa gcaggcccag gacauacaac ccgcuaugca ggcuucaugg 3960  
cccaaagugg aacaauuuug ggccagacac auguggaacu ucauuagcgg cauccauuac 4020  
cucgcaggau ugucaacacu gccagggaac cccgcggugg cuuccaugau ggcauucagu 4080  
gccgcccuca ccaguccguu gucgaccagu accaccaucc uucucaacau caugggaggc 4140  
ugguuagcgu cccagaucgc accaccgcg ggggccaccg gcuuugucgu caguggccug 4200  
gugggggucg ccgugggcag cauaggccug gguaaggugc ugguaggacau ccuggcagga 4260  
uauggugcgg gcauuucggg ggcccucguc gcauucaaga ucaugucugg cgagaagccc 4320  
ucuauggaag augucaucaa ucuacugccu gggauccugu cuccgggagc ccugguggug 4380  
ggggucaucu gcgcggccau ucugcgccgc cacgugggac cgggggaggg cgcgguccaa 4440  
uggaugaaca ggcuuuugc cuuugcuucc agaggaaacc acgucgcccc uacucacuac 4500  
gugacggagu cggaugcguc gcagcgugug acccaacuac uuggcucucu uacuauaacc 4560  
agccuacuca gaagacucca caauuggaua acugaggacu gcccacucc augcuccgga 4620  
uccuggcucc gcgacgugug ggacuggguu ugcaccaucu ugacagacuu caaaaauugg 4680  
cugaccucua aauuguuccc caagcugccc ggccucccu ucaucucuug ucaaaagggg 4740  
uacaagggug ugugggccgg cacuggcauc augaccacgc gcugcccuug cggcgccaac 4800

aucucuggca auguccgccu gggcucuaug aggaucacag ggccuaaaac cugcaugaac 4860  
accuggcagg ggaccuuucc uaucaauugc uacacggagg gccagugcgc gccgaaaccc 4920  
cccacgaacu acaagaccgc caucuggagg guggcggccu cggaguacgc ggaggugacg 4980  
cagcaugggu cguacuccua uguaacagga cugaccacug acaaucugaa aaauccuugc 5040  
caacuaccuu cuccagaguu uuucuccugg guggacggug ugcagaucca uagguuugca 5100  
cccacaccaa agccguuuuu ccgggaugag gucucguucu gcguugggcu uaauuccuau 5160  
gcugucgggu cccagcuucc cugugaaccu gagcccgacg cagacguauu gagguccaug 5220  
cuaacagauc cgccccacau cacggcggag acugcggcgc ggcgcuuggc acggggauca 5280  
ccuccaucug aggcgagcuc cucagugagc cagcuauag caccgucgcu gcgggccacc 5340  
ugcaccaccc acagcaacac cuaugacgug gacauggucg augccaaccu gcucauggag 5400  
ggcggugugg cucagacaga gccugagucc agggugcccg uucuggacuu ucucgagcca 5460  
auggccgagg aagagagcga ccuugagccc ucaauaccu cggagugcau gcucuccagg 5520  
agcggguuuc cacgggccuu accggcuugg gcacggccug acuacaaccc gccgcucgug 5580  
gaaucgugga ggaggccaga uuaccaaccg cccaccguug cugguugugc ucucgggggg 5640  
cccaagaagg ccccgacgcc uccccaaagg agacgccgga cagugggucu gagcgagagc 5700  
accauaucag aagcccucca gcaacuggcc aucaagaccu uggccagcc ccccucgagc 5760  
ggugaugcag gcucguccac gggggcgggc gccgccgaau ccggcggucc gacgucccu 5820  
ggugagccgg ccccccucaga gacagguucc gccuccucua ugccccccu cgaggggggag 5880  
ccuggagauc cggaccugga gucugaucag guagagcuuc aaccucgggg ccaggggggg 5940  
gggguagcuc ccgguucggg cucggggucu uggucuacuu gcuccgagga ggacgauacc 6000  
accgugugcu gcuccauguc auacuccugg accggggcuc uaauaacucc cuguagcccc 6060  
gaagaggaaa aguugccaau caaccuuug aguaacucgc uguugcgaua ccauaacaag 6120  
guguacugua caacaucaaa gagcgccuca cagagggcua aaaagguaac uuuugacagg 6180  
acgcaagugc ucgacgcca uuaugacuca gucuuaaagg acaucaagcu agcggcuucc 6240  
aaggucagcg caaggcuccu caccuuggag gaggcgugcc aguugacucc accccauucu 6300  
gcaagaucca aguauggauu cggggccaag gagguccgca gcuuguccgg gagggccguu 6360  
aaccacauca aguccgugug gaaggaccuc cuggaagacc caciaaacacc aaaucccaca 6420  
accaucaugg ccaaaaauga gguguucugc guggaccccg ccaagggggg uaagaaacca 6480  
gcucgccuca ucguuuaccc ugaccucggc guccgggucu gcgagaaaau ggcccucua 6540

gacauuacac aaaagcuucc ucaggcggua augggagcuu ccuauggcuu ccaguacucc 6600  
ccugcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagga ccccaugggg 6660  
uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720  
gaguccauau accaggccug cucccugccc gagggaggccc gcacugccau acacucgcug 6780  
acugagagac uuuacguagg agggcccaug uucaacagca agggucacac cuvcgguuac 6840  
agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccu cacaugcuau 6900  
gugaaagccc uagcggccug caaggcugcg gggauaguug cgcccacaau cucagaaagc 6960  
caggggacug aggaggacga gcggaaccug agagccuua cggaggccau gaccagguac 7020  
ucugccccuc cuggugaucc ccccagaccg gaauaugacc uggagcuau acauccugu 7080  
uccucaaauug ugucuguggc guugggccc cggggccgcc gcagauacua ccugaccaga 7140  
gacccaacca cuccacucgc ccgggcugcc ugggaaacag uuagacacuc ccuaucaau 7200  
ucauggcugg gaaacaucau ccaguaugcu ccaaccuau ggguucgcau gguccuaau 7260  
acacacuucu ucuccauucu caugguccaa gacaccucgg accagaaccu caacuugag 7320  
auguauggau caguauacuc cgugaauccu uggaccuuc cagccauau ugagagguua 7380  
cacgggcuug acgccuuuuc uaugcacaca uacucucacc acgaacugac gcggguggc 7440  
ucagcccua gaaaacuugg ggcgccacc cucagggugu ggaagagucg ggcucgcgc 7500  
gucagggcgu ccucaucuc ccguggagg aaagcggccg uuugcggccg auaucucuuc 7560  
aaugggcgg ugaagaccaa gcucuaacuc acuccauugc cggaggcgc ccuacuggac 7620  
uuauccaguu gguuaccgu cggcgccggc gggggcgaca uuuuucacag cgugucgcgc 7680  
gcccgaaccc gcucuuacu cuucggccua cuccuacuu ucguaggggu aggcucucuc 7740  
cuacucuccg cucgguagag cggcacacac uagguacacu ccuagcuua cuguuccuu 7800  
uuuuuuuuuu uuuuuuuuu uuuuuuuuu uuuuuuuuu cuuuuuuuuu uuuuuuccuc 7860  
uuucuuuccu ucuaucua uucuaucuu uuucuggug gcuccauuu agccuaguc 7920  
acggcuagcu gugaaagguc cgugagccgc augacugcag agagugccgu aacuggucuc 7980  
ucugcagauc augu 7994

&lt;210&gt; 9

&lt;211&gt; 340

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic RNA

<400> 9

```
accugccccc aaauaggggcg acacuccgcc augaaucacu cccugugag gaacuacugu 60
cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120
ccccucccg ggagagccau aguggucugc ggaaccggug aguacaccgg aaaugccggg 180
aagacugggu ccuuucuugg auaaaccac ucuaugcccg gccauuuggg cgugcccccg 240
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300
cgcuugcgag ugccccggga ggucucguag accgugcacc 340
```

<210> 10

<211> 340

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic RNA

<400> 10

```
acccgcccc aaauaggggcg acacuccgcc augaaucacu cccugugag gaacuacugu 60
cuucacgcag aaagcgucua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120
ccccucccg ggagagccau aguggucugc ggaaccggug aguacaccgg aaaugccggg 180
aagacugggu ccuuucuugg auaaaccac ucuaugcccg gccauuuggg cgugcccccg 240
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300
```

ugcuugcgag ugccccggga ggucucguag accgugcacc

340

&lt;210&gt; 11

&lt;211&gt; 236

&lt;212&gt; RNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: synthetic RNA

&lt;400&gt; 11

agcggcacac acuagguaca cuccauagcu aacuguuccu uuuuuuuuuu uuuuuuuuuu 60  
uuuuuuuuuu uuuuuuuuuu uuuuuuuuuu uuuuuuuucc uuuuuuuucc cuucucaucu 120  
uauucuacuu uuuuuuuugg ugguccauuc uuagcccuag ucacggcuag cugugaaagg 180  
uccgugagcc gcaugacugc agagagugcc guaacugguc ucucugcaga ucaugu 236

&lt;210&gt; 12

&lt;211&gt; 232

&lt;212&gt; RNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: synthetic RNA

&lt;400&gt; 12

agcggcacac auuagcuaca cuccauagcu aacuguuccu uuuuuuuuuu uuuuuuuuuu 60  
uuuuuuuuuu uuuuuuuuuu uuuuuuuuuu uuucccucu uuuucccuuc ucaucuauu 120

cuacuuucuu ucuugguggc uccaucuuag ccugggucac ggcuagcugu gaaagguccg 180  
ugagccgcau gacugcagag agugccguaa cuggucucuc ugcagaucau gu 232

<210> 13

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 13

cgggagagcc atagtgg

17

<210> 14

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 14

agtaccacaa ggcctttcg

19

<210> 15

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 15

ctgcggaacc ggtgagtaca c

21

<210> 16

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 16

aacaagatgg attgcacgca

20

<210> 17

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA



&lt;400&gt; 17

cgtcaagaag gcgatagaag

20

&lt;210&gt; 18

&lt;211&gt; 30

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: synthetic DNA

&lt;400&gt; 18

gcactctctg cagtcatgcg gctcacggac

30

&lt;210&gt; 19

&lt;211&gt; 28

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: synthetic DNA

&lt;400&gt; 19

cccctgtgag gaactactgt cttcacgc

28

<210> 20

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 20

ccgggagagc catagtggtc tgcg

24

<210> 21

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 21

ccactcaaag aaaaagtgtg acgagctcgc

30

<210> 22

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 22

ggcttgggca cggcctga

18

<210> 23

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 23

gcggtgaaga ccaagctcaa actcactcca

30

<210> 24

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 24

agaacctgcg tgcaatccat c

21

<210> 25

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 25

cccgtcatga gggcgtcggg ggc

23

<210> 26

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 26

accagcaacg gtgggcgggt ggtaatc

27

<210> 27

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 27

ggcacgcgac acgctgtg

18

<210> 28

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 28

agctagccgt gactagggct aagatggagc

30

#### 【配列表フリーテキスト】

配列番号 1、2、7 及び 8 で示される配列は、レプリコンである。

配列番号 9 ～ 12 で示される配列は合成 RNA である。

配列番号 13 ～ 28 で示される配列は、合成 DNA である。

#### 【図面の簡単な説明】

##### 【図 1】

図 1 は、本発明に係る HCV-RNA レプリコンを作製するための鋳型 DNA の構築手順を示す概略図である。図 1 の上段は、pJFH1 及び pJCH1 のウイルスゲノム挿入

領域の構造を示す。図1の下段は、pJFH1及びpJCH1のウイルスゲノム挿入領域の一部をネオマイシン耐性遺伝子とEMCV IRESを含むDNA断片で置換することにより構築したプラスミドDNA pSGREP-JFH1及びpSGREP-JCH1のウイルスゲノム挿入領域の構造を示す。図中の記号は以下のとおりである。T7: T7 RNAプロモーター、G: 挿入したJFH-1又はJCH-1由来DNAの5'端の上流かつT7 RNAプロモーター配列の3'端の下流に挿入したdGTP、5'NTR: 5'非翻訳領域、Core: コアタンパク質、3'NTR: 3'非翻訳領域。E1、E2: エンベロープタンパク質。NS2、NS3、NS4A、NS4B、NS5A、NS5B: 非構造タンパク質。Age I、Cla I、Xba I: 制限酵素Age I、Cla I及びXba Iの切断部位。GDD: NS5Bタンパク質の活性中心に相当するアミノ酸モチーフGDDの位置。neo: ネオマイシン耐性遺伝子、EMCV IRES: EMCV IRES (脳心筋炎ウイルスの内部リボソーム結合部位)。

**【図2A】**

図2は、rSGREP-JFH1の塩基配列を示す。

**【図2B】**

図2は、rSGREP-JFH1の塩基配列を示す。

**【図2C】**

図2は、rSGREP-JFH1の塩基配列を示す。

**【図2D】**

図2は、rSGREP-JFH1の塩基配列を示す。

**【図2E】**

図2は、rSGREP-JFH1の塩基配列を示す。

**【図2F】**

図2は、rSGREP-JFH1の塩基配列を示す。

**【図3A】**

図3は、rSGREP-JCH1の塩基配列を示す。

**【図3B】**

図3は、rSGREP-JCH1の塩基配列を示す。

**【図3C】**

図 3 は、rSGREP-JCH1 の塩基配列を示す。

【図 3 D】

図 3 は、rSGREP-JCH1 の塩基配列を示す。

【図 3 E】

図 3 は、rSGREP-JCH1 の塩基配列を示す。

【図 3 F】

図 3 は、rSGREP-JCH1 の塩基配列を示す。

【図 4】

図 4 は、rSGREP-JFH1、rSGREP-JFH1/GND 及び rSGREP-JFH1/dGDD をそれぞれトランスフェクションした Huh7 細胞のコロニー形成を示した写真である。トランスフェクションした RNA の量は、上段の 3 つはいずれも 100 ng、下段の 3 つはいずれも 300 ng である。

【図 5】

図 5 は、rSGREP-JFH1 及び rSGREP-JCH1 をそれぞれトランスフェクションした Huh7 細胞の、培地中の G418 の濃度が 0.5mg/ml である場合のコロニー形成を示した写真である。トランスフェクションした RNA の量は、いずれも 100 ng である。

【図 6】

図 6 は、トランスフェクション細胞のコロニー形成能に対する、Mung Bean Nuclease 処理の影響を示した写真である。トランスフェクションした rSGREP-JFH1 RNA の量は、いずれも 100 ng である。いずれも培地中の G418 の濃度は 1.0mg/ml である。

【図 7】

図 7 は、rSGREP-JFH1 をトランスフェクションして樹立したレプリコン複製細胞クローン由来のトータル細胞性 RNA を新たな Huh7 細胞に再トランスフェクションした場合に示されるコロニー形成を示す写真である。左側写真は、レプリコン複製細胞クローン 6 由来のトータル細胞性 RNA を用いて 96 コロニーの形成がみとめられた結果、右側写真：プールクローン由来のトータル細胞性 RNA を用いて 77 コロニーの形成がみとめられた結果を示す。いずれもレプリコン RNA を  $1 \times 10^7$  コピー含む量を再トランスフェクションした。

## 【図 8】

図 8 は、rSGREP-JFH1 をトランスフェクションして樹立したレプリコン複製細胞クローン由来のトータル細胞性 RNA を新たな Huh7 細胞に再トランスフェクションして得た細胞クローン由来のトータル RNA に対して、rSGREP-JFH1 特異的プローブを用いてノーザンブロット法による検出を行った結果を示す写真である。レーンの説明は以下のとおりである。10<sup>8</sup>: Huh7 細胞から抽出したトータル RNA に試験管内で合成したレプリコン RNA を 10 の 8 乗コピー加えたサンプル、10<sup>7</sup>: Huh7 細胞から抽出したトータル RNA に試験管内で合成したレプリコン RNA を 10 の 7 乗コピー加えたサンプル、Huh7: トランスフェクションしていない Huh7 細胞から抽出したトータル RNA、プールクローン: プールクローンから抽出したトータル RNA、1~11: 細胞クローン 1~11 のそれぞれから抽出したトータル RNA。「レプリコン RNA」は、rSGREP-JFH1 の分子量サイズを示すマーカー、「28S」は 4.5kb の分子量サイズを示すリボソーム RNA マーカー、「18S」は 1.9kb の分子量サイズを示すリボソーム RNA マーカーの泳動位置を示す。

## 【図 9】

図 9 は、rSGREP-JFH1 又は rSGREP-JCH1 由来複製レプリコン RNA が再トランスフェクションされた細胞クローンにおける、ネオマイシン耐性遺伝子の宿主細胞のゲノム DNA への組み込みの有無を示す写真である。左側写真のレーンの説明は以下のとおりである。M: DNA 分子量マーカー、1~8: rSGREP-JFH1 由来細胞クローン 1~8、N: トランスフェクションしていない Huh7 細胞、P: 陽性対照（ネオマイシン耐性遺伝子の PCR 増幅産物）。一方、右側写真のレーンの説明は以下のとおりである。M: DNA 分子量マーカー、1~6: rSGREP-JCH1 由来細胞クローン 1~6。

## 【図 10】

図 10 は、rSGREP-JFH1 又は rSGREP-JCH1 由来複製レプリコン RNA が再トランスフェクションされた細胞クローンにおいて、発現された NS3 タンパク質の検出結果を示す写真である。左側写真のレーン 1~8: rSGREP-JFH1 由来細胞クローン 1~8。右側写真のレーン 1~6: rSGREP-JCH1 由来細胞クローン 1~6。右



側写真のレーンの P：NS3タンパク質（陽性対照）、N：トランスフェクションしていないHuh7細胞から抽出したタンパク質（陰性対照）。

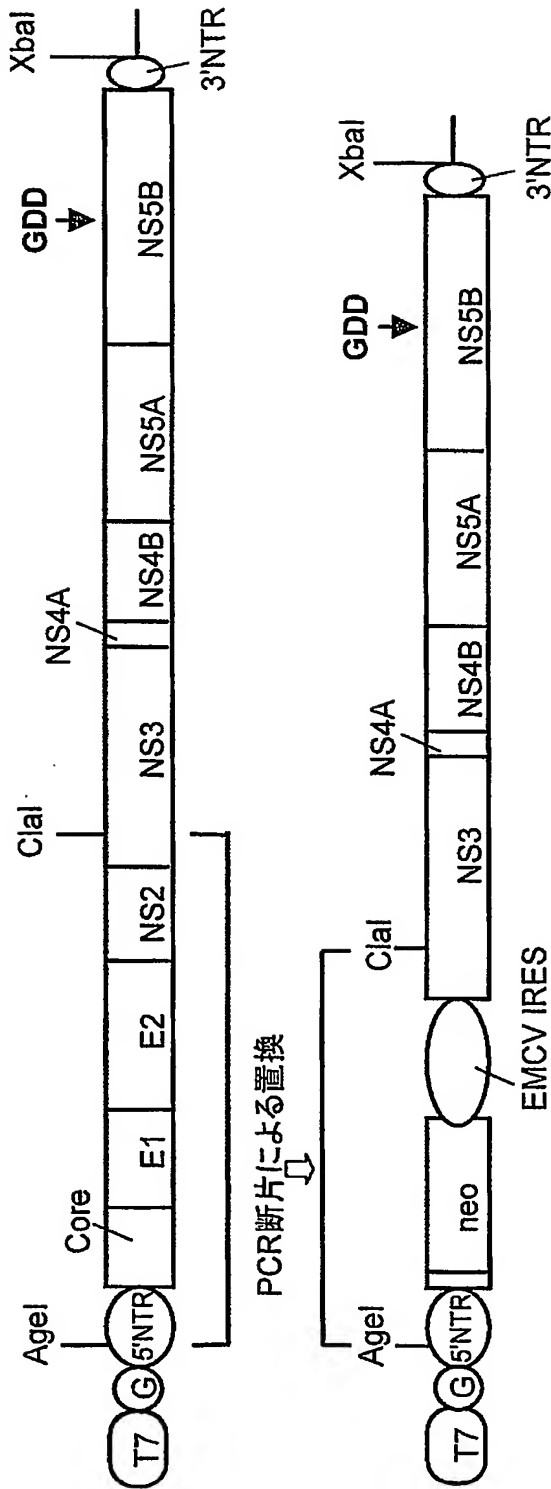
【図 11】

図 11 は、rSGREP-JFH1由来の複製レプリコンRNAのHuh7細胞への再トランスフェクションを経て樹立した21の細胞クローンから取得したレプリコンRNA中の塩基変異の位置を示す図である。変異の位置は、表4に示す塩基番号を付記した縦線で示している。太い縦線は非同義置換、細い縦線は同義置換を表す。

【書類名】 図面

【図 1】

pJFH1, pJCH1



pSGREP-JFH1, pSGREP-JCH1

【図 2 A】

10	20	30	40	50	60
ACCUGCCCG	AAUAGGGCG	ACACUCCGC	AUGAUCACU	CCCCUGAG	GAACUACUGU
70	80	90	100	110	120
CUUACGCG	AAAGCGCUA	GCCAUGGCG	UAGUAUGAGU	GUUGUACAGC	CUCCAGGCC
130	140	150	160	170	180
CCCCUCCCG	GGAGAGCCAU	AGUGGUCUGC	GGAACCGGCG	AGUACACCGG	AAUUGCCGGG
190	200	210	220	230	240
AAGACUGGCU	CCUUUCUGG	AUAAACCCAC	UCUAUGCCCG	GCCAUUUGGG	CGUGCCCCCG
250	260	270	280	290	300
CAAGACUGCU	AGCCGAGUAG	CGUUGGGUUG	CGAAGGCCCU	UGUGGUAUCG	CCUGAUAGGG
310	320	330	340	350	360
CGCUUGCGAG	UGCCCCGGGA	GGUCUUGUAG	ACCGGCGACC	AUGAGCAAA	AUCCUAAACC
370	380	390	400	410	420
UCAAGAGAA	ACCAAAAGAA	ACACCAACCG	UCGCCCAUG	AUUGAACAA	AUGGAUUGCA
430	440	450	460	470	480
CGCAGGUUCU	CGGCGGCUU	GGGUGGAGAG	GCUAUUGGG	UAUGACUGGG	CACACAGAGC
490	500	510	520	530	540
AAUUGGCGCC	UCUUAUGCG	CGUGGUGCG	GCUGUCAGCG	CAGGGGCGCC	CGGUUCUUCU
550	560	570	580	590	600
UGUCAGAGCC	GACCGUGCG	GGGCCUGAA	UGAACUGCAG	GAGAGGCGAG	CGCGGCUAUC
610	620	630	640	650	660
GUGGCGGCG	ACGACGGGCG	UCCCUUGGCG	AGCUGUGCUC	GACGUUGUCA	CGAAGCGGG
670	680	690	700	710	720
AAGGGACUGG	CUUCUAUUG	GCGAAGUGCC	GGGCGAGGAU	CUCCUGCAU	CUACCCUGGC
730	740	750	760	770	780
UCCUGGCGAG	AAAGUAUCCA	UCAUGGCUGA	UGCAAUGCGG	CGCUGCAUA	CGCUGAUCC
790	800	810	820	830	840
GGCUUCCUGC	CCAUUGGACC	ACCAAGCGAA	ACAUGGCAUC	GAGCGAGCAC	GUACUUGGAU
850	860	870	880	890	900
GGAAGCGGCU	CUUGUGCAUC	AGGAUGAUCU	GGACGAGGAG	CAUCAGGGGC	UUGGCCCGGC
910	920	930	940	950	960
CGAACUGUUC	GCCAGGCUCA	AGGCGGCGAU	GCCCCAGGCG	GAGGAUUCUG	UCGUGACCCA
970	980	990	1000	1010	1020
UGGCGAGGCC	UGCUUGCGA	AUAUCAUGGU	GGAAAUUGGC	CGCUUUCUG	GAUUCAUCCA
1030	1040	1050	1060	1070	1080
CUGUGGCGCG	CGGGUGUGG	CGGACCGCUA	UCAGGACAU	GCGUUGGCUA	CCCGGGAUUA
1090	1100	1110	1120	1130	1140
UGCUGAGGAG	CUUGGCGGCG	AAUGGGCUGA	CCGCUUCCUC	GUGCUUACG	GUUUGCGCGC
1150	1160	1170	1180	1190	1200
UCCCGAUUCG	CAGCGCAUOG	CCUUCUAUOG	CCUUCUUGAC	GAGUUCUUCU	GAGUUUAAC
1210	1220	1230	1240	1250	1260
CCUUCUCCUC	CCCCCCCCU	AAAGUUAUCG	GCGGAGCGCG	CUUGGAUUA	GGCGGUGUG
1270	1280	1290	1300	1310	1320
CGUUGUGCUA	UAUGUAUUAU	UCCACCAUUA	UGCGUUCUUA	UGGCAUUGUG	AGGGCCCGGA
1330	1340	1350	1360	1370	1380
AACCGGCGCC	CGUUCUUCUG	ACGAGCAUUC	CUAGGGGUCU	UUCGCCUUC	GCCAAAGGAA

【図2B】

1390	1400	1410	1420	1430	1440
UGCAAGGUCU	GUUGAAUGUC	GUGAAGGAAG	CAGUUCUCU	GGAAGCUCU	UGAAGACAAA
1450	1460	1470	1480	1490	1500
CAACGUCUGU	AGCGACCCUU	UGCAGGCAGC	GGAACCCCCC	ACCUGGCGAC	AGGUGCCUCU
1510	1520	1530	1540	1550	1560
GCGGCCAAAA	GCCACUGUA	UAAGAUAAC	CUGCAAAAGC	GGCACAACCC	CAGUGOCACG
1570	1580	1590	1600	1610	1620
UUGUGAGUUG	GAUAGUUGUG	GAAGAGUCA	AAUGGCUCUC	CUCAGCGUA	UUCAAACAAGG
1630	1640	1650	1660	1670	1680
GCGUGAGGGA	UGCCAGAGAG	GUACCCCAU	GUAGGGGUA	UGAUCUGGG	CCUGCGUGCA
1690	1700	1710	1720	1730	1740
CAUGCUUUAC	AUGUGUUUAG	UCCAGGUUAA	AAAAACGUCU	AGGCCCCCGG	AACCACGGGG
1750	1760	1770	1780	1790	1800
ACUGCUUUU	CCUUGAABAA	ACAAGAUAGU	ACCAUGGCUC	CCUACACUGC	UUUAGCCACG
1810	1820	1830	1840	1850	1860
CAAAACAGAG	GCCUCCUGGG	CGCCAUAGUG	GUGAGUAUCA	CGGGGCGUGA	CAGGACAGAA
1870	1880	1890	1900	1910	1920
CAGGCCGGGG	AGUCCAAAU	CCUGUCCACA	GUUCUCACGU	CCUUCUCUGG	AACAACCAUC
1930	1940	1950	1960	1970	1980
UCGGGGGUUU	UGUGGACUGU	UUACCAACGA	CGUGGCAACA	AGACUUCAGC	CGGCUUACGG
1990	2000	2010	2020	2030	2040
GGUCCGUGCA	CGCAGAUUA	CUCCAGUGCU	GAGGGGACU	UGGUAAGCUG	GCCAGCCCC
2050	2060	2070	2080	2090	2100
CCUGGGACCA	AGUCUUUGGA	GCCGUGCAAG	UGUGGAGCCG	UCGACCUAUA	UCUGGUCACG
2110	2120	2130	2140	2150	2160
CGGAACGUCG	AUGUCAUCCC	GGCUCCGAGA	CGCGGGGACA	AGCGGGGAGC	AUUGCCUCCC
2170	2180	2190	2200	2210	2220
CCGAGACCCA	UUUCGACCUU	GAAGGGGUCC	UCCGGGGGGC	CGGUGUCUUG	CCCUAGGGGC
2230	2240	2250	2260	2270	2280
CACGUCUGUG	GGCUCUUCCG	AGCAGCUGUG	UGUCUUCGGG	GCGUGGCCAA	AUCCAUCCAU
2290	2300	2310	2320	2330	2340
UUCAUCCCCG	UUGAGACACU	CGACGUUGUU	ACAAGGCCCC	CCACUUCAG	CGACACACAG
2350	2360	2370	2380	2390	2400
ACGCCACCCG	CUGGCCCCCA	GACCUAUCAG	GUCCGGUACU	UGCAUGCUCC	AACUGGCAGU
2410	2420	2430	2440	2450	2460
GGAAAGAGCA	CCAAGGUCCC	UGUCGCGUAU	GCCGCCCCAG	GGUACAAAGU	ACUAGUGCUU
2470	2480	2490	2500	2510	2520
AACCCCUUGG	UAGCUGGCAC	CCUGGGGUUU	GGGGGUAACC	UAUCCAAAGC	ACAUGGCAUC
2530	2540	2550	2560	2570	2580
AAUCCCAACA	UUGGACUUGG	AGUCAGGACC	GUGAUGACCG	GGGAGGCCAU	CAGGUACUCC
2590	2600	2610	2620	2630	2640
ACAUAGGGCA	AAUUCUUGC	CGAUGGGGGC	UGCGUAGCG	GCGCCUAUGA	CAUCACAUUA
2650	2660	2670	2680	2690	2700
UGCGAUGAAU	GCCACGCUUG	GGUUGCUACC	UCCAUCUUCG	GCAUCCGAAC	GGUCCUUGAU
2710	2720	2730	2740	2750	2760
CAAGCAGAGA	CAGCCGGGGU	CAGACUAACU	GUGCUUGCUA	CGCCACACCC	CCCCGGGUCA

【図 2 C】

2770	2780	2790	2800	2810	2820
GUAGCAAGCC	CCCAUCCGA	UAUAGAGAG	GUAGGCCUG	GGGGGAGGG	UGAGAUCCCC
2830	2840	2850	2860	2870	2880
UUCUAUGGGA	GGGCGAUUC	CCUAUCCUG	AUCAAGGGAG	GGAGACACU	GAUUAUCCGC
2890	2900	2910	2920	2930	2940
CACUCAAGA	AAAGUGUGA	CGAGCUCGC	GCGGCCUUC	GGGCAUUGG	CUUGAAUCCC
2950	2960	2970	2980	2990	3000
GUGGCAUAC	AUAGAGGGU	GGACGUCUC	AUAUAACAG	CUCAGGAGA	UGUGGUGGUC
3010	3020	3030	3040	3050	3060
GUCCCAACG	ACGCCUCAU	GACGGGUAC	ACUGGAGACU	UUGACTUCGU	GAUGGACTGC
3070	3080	3090	3100	3110	3120
AUUGUAGGG	UCACCAAGC	UGUCCACUC	AGCCUGGAC	CCACCUCCAC	UAUAACACA
3130	3140	3150	3160	3170	3180
CAGACUGUC	CACAAGACG	UGUCUACGC	AGUCAGCGC	GCGGGCGAC	AGGUAGAGGA
3190	3200	3210	3220	3230	3240
AGACAGGGA	CUUAUAGUA	UGUUCCACU	GGUGAACAG	CCUCAGGAU	GUUUGACAGU
3250	3260	3270	3280	3290	3300
GUAGUGCUU	GUGAGUGCU	CGACGAGGG	GCUUGUGGU	ACGAUCCAC	ACCAGCGGAG
3310	3320	3330	3340	3350	3360
ACCAACGUA	GGCUUAGAG	GUUUUACAC	ACGCCCGGC	UACCGGUGG	UCAAGACCAU
3370	3380	3390	3400	3410	3420
CUUGAAUUU	GGGAGGAGU	UUUACCCGC	CUACACACA	UAGAGGCCA	CUUCCUCUCC
3430	3440	3450	3460	3470	3480
CAAACAAAG	AAGCGGGGA	GAACUUCCG	UACCUAGUAG	CCUACCAAG	UACGGUGUGC
3490	3500	3510	3520	3530	3540
GCCAGAGCA	AGGCCCUCC	CCGUCUUGG	GACGCCAGU	GGAGUGCCU	GGCCCGACCC
3550	3560	3570	3580	3590	3600
AAGCCUACG	UUGCGGCCC	CACACUUCU	CUGUACCGU	UGGGCCUUA	UACCAUGAG
3610	3620	3630	3640	3650	3660
GUACCCUCA	CACACCCUG	GACGAAGUAC	AUCGCCACU	GCAUGCAAG	UGAACCUAG
3670	3680	3690	3700	3710	3720
GUCAAGACA	GCAUGUGGU	CCUAGUGGA	GGAGCCUGG	CAGCCGUGC	CGCAUAUUGC
3730	3740	3750	3760	3770	3780
CUGGGACCG	GAUGGUGUC	CAUCAUCCG	CGCUUGCAG	UCABCCAGG	AGUUGUCGUU
3790	3800	3810	3820	3830	3840
GCGCCGUAU	AGGAGGUCC	GUUUGAGGU	UUUGAUAGA	UGGAGGAUG	CGCCUUGAG
3850	3860	3870	3880	3890	3900
GCGGCUUCA	UCCAGAGGG	GCAGCGGUA	GCCGAGAGU	UGAAGUCCAA	GAUCCAGGC
3910	3920	3930	3940	3950	3960
UUGCUGCAG	AGGCCUUAU	GCAGGCCAG	GACAAACAC	CCGCCUAGCA	GGCUUACUGG
3970	3980	3990	4000	4010	4020
CCCAAGGUG	AACAAUUGU	GGCAGACAC	AUGUGGACU	UCAUAGCCG	CAUCCAAUAC
4030	4040	4050	4060	4070	4080
CCCGCAGGA	UGUCAACACU	CCAGGGAAC	CCCGCGUGG	CUUCCAUAGU	GGCAUUCAGU
4090	4100	4110	4120	4130	4140
GCGGCCUCA	CCAGUUCGU	GUCCACAGU	ACCACCCUCC	UUUCACACU	CAUGGGAGGC

【図 2 D】

4150	4160	4170	4180	4190	4200
UGGUUAGCGU	CCAGAUCCG	ACCACCCGCG	GGGCCCAACG	GCUUUGUGU	CAGUGGCCUG
4210	4220	4230	4240	4250	4260
GGGGGGGCTG	CCGUGGCCAG	CAUAGGCCUG	GGUAGGUGGC	UGGUGGACAU	CCUGGCAGGA
4270	4280	4290	4300	4310	4320
UAUGGUGCGG	GCAUUGCGGG	GGCCUUCGUC	GCAUUCAGA	UCAUGUCUGG	CGAGAAAGGCC
4330	4340	4350	4360	4370	4380
UCUAGGAGAG	AUGUCAUCAA	UCUUCUGCCU	GGGAUCCUGU	CUCCCGGAGC	CCUGGUGGUG
4390	4400	4410	4420	4430	4440
GGGGUUAUCU	GCGCGGCCAU	UCUGCGCGGC	CACGUGGGAC	CGGGGGAGGG	CGCGGUCCAA
4450	4460	4470	4480	4490	4500
UGGAUGRACA	GGCUUUAUUG	CUUUGCUUCC	AGAGGAAACC	ACGUGGCCCC	UACUCACUAC
4510	4520	4530	4540	4550	4560
GUGACGGAGU	CGGAUGCGUC	GCAGCGUGUG	ACCCAAACUAC	UUGGCUUCU	UACUAAUACC
4570	4580	4590	4600	4610	4620
AGCCUACUCA	GAAGACUCCA	CAAUUGGAUA	ACUGAGGACU	GGCCAUCCC	AUGCUCCGGA
4630	4640	4650	4660	4670	4680
UCCUGGCUCC	GCGACGUGUG	GGACUGGGUU	UGCAACAUCC	UGACAGACUU	CAAAAUUUGG
4690	4700	4710	4720	4730	4740
CGAGCCUUA	AAUUGUCCC	CAAGCGGCC	GGCCUCCCU	UCAUCUUCUG	UCAAAAGGGG
4750	4760	4770	4780	4790	4800
UACAGGGUG	UGUGGGCCGG	CACUGGCACU	AUGACCAACG	CGUGCCUUG	CGCGGCCAAC
4810	4820	4830	4840	4850	4860
AUCUUGGCA	AUGUCCGCCU	GGGCUCCUUG	AGGAUACACG	GGCCUAAAC	CUGCAUGAAC
4870	4880	4890	4900	4910	4920
ACCGGCGAGG	GGACCUUUC	UACCAAUUGC	UACACGGAGG	GCCAGUGGCG	GGCGAAACCC
4930	4940	4950	4960	4970	4980
CCCAAGAACU	ACAGGACCGC	CAUCUGGAGG	GUGGCGGCCU	CGAGUAACGC	GGAGGUGAGC
4990	5000	5010	5020	5030	5040
CAGCAUGGGU	GGUACUCCUA	UGUAACAGGA	CUGACCACUG	ACAUCUGAA	AAUUCUUGC
5050	5060	5070	5080	5090	5100
CAACUACCUU	CUCCAGAGUU	UUUCUCCUGG	GUGGAAGGUG	UGCAGAUCCA	UAGGUUUGCA
5110	5120	5130	5140	5150	5160
CCCAACCCAA	AGCGGUUUUU	CCGGGAGGAG	GGCUUGUUCU	GGUUGGGGU	UAAUUCUUAU
5170	5180	5190	5200	5210	5220
GCUUGUGGGU	CCCAGCUCCC	CUUGGAAACU	GAGCCGAGCG	CAGAGUAUUU	GAGGUCCATG
5230	5240	5250	5260	5270	5280
CUAACAGAUU	CGCCCAACAU	CACGGGGGAG	ACCGCGGGGC	GGCGCUUGGC	ACGGGGAGCA
5290	5300	5310	5320	5330	5340
CCUCCAUUCG	AGGGAGCCUC	CUCAUGGAGC	CAGCUUUCAG	CACCGUUGCU	GCGGGCCACC
5350	5360	5370	5380	5390	5400
UGCAACACCC	ACAGCAACAC	CUAUGACGUG	GACAUGGUUG	AUGCCAAACU	GCUCAUGGAG
5410	5420	5430	5440	5450	5460
GGGGGUGUGG	CUCAGACAGA	GCUGAGUCC	AGGGUGCCCG	UUUUGGACUU	UCUGAGGCCA
5470	5480	5490	5500	5510	5520
AUUGCCGAGG	AGGAGAGCGA	CCUUGAGGCC	UCAAUACCAU	CGGAGUGCAU	GCUCUCCAGG

【図2E】

5530 5540 5550 5560 5570 5580  
AGCGGGUUC CAGGGGCCUU ACCGGCUUG GCAAGGCUU ACUACAAACC GCGGCUUGG  
5590 5600 5610 5620 5630 5640  
GAAUCUGGA GGAGGCCAG UUAACCAACG CCCACCGUUG CUGGUGUGG UCUCCCCC  
5650 5660 5670 5680 5690 5700  
CCCAAGAGG CCGGACGCC UCCCCAAGG AGACGCGGA CAGUGGUCU GAGCGAGAGC  
5710 5720 5730 5740 5750 5760  
ACCAUAUCAG AAGCCUCCA GCAACUGGC AUCAAGACCU UUGGCGAGC CCCCUGAGC  
5770 5780 5790 5800 5810 5820  
GGUGAUGAG GCUUGUCCAC GGGGGCGGC GCGGCGAAG CCGGCGUCC GAGUGCCCCU  
5830 5840 5850 5860 5870 5880  
GGUGAGCGG CCCCCAGAG GACAGGUCC GCCUCCUUA UGCCCCCUU CGAGGGGAG  
5890 5900 5910 5920 5930 5940  
CCUGGAGAG CCGACCUGA GUCUGAUCAG GUAAGACUUC AACUCCCC CCAGGGGGG  
5950 5960 5970 5980 5990 6000  
CGGUAUGUC CCGUUGGG CUGGGGUCU UGGUUAUCU GCUCCAGGA GGAAGUACC  
6010 6020 6030 6040 6050 6060  
ACCGUGUGU GCUCCAGUC AUACUCCUG ACCGGGCUU UAAUAACUC CUGUAGCCCC  
6070 6080 6090 6100 6110 6120  
GAAAGGAGA AGUUGCAAU CAACCUUUG AGUAACUCC UGUUGGUAU CCADAACAG  
6130 6140 6150 6160 6170 6180  
GCGUACUGA CAACAUCAA GAGGCGCCA CAGAGGCUA AAAAGGUAU UUUUACAGG  
6190 6200 6210 6220 6230 6240  
ACGCAAGUC UCGACGCCA UUAUGACUA GUCUUAAGG ACAUCAGCU AGCGGCUUC  
6250 6260 6270 6280 6290 6300  
AAGUACAGC CAAGGCUCC CACCUUGAG GAGGCGUCC AGUUGACUC ACCCAUUCU  
6310 6320 6330 6340 6350 6360  
GCAAGAUCA AGUAGGAGU CCGGGCCAG GAGGUCGCA GCUUGGCGG GAGGGCGGU  
6370 6380 6390 6400 6410 6420  
AAACAUAUA AGUCCGUGU GAAGGACUC CUGGAAGAC CACAACACC AAUUCCTCA  
6430 6440 6450 6460 6470 6480  
ACCAUACAG CCAAAAGUA GUGUUCUGC GUAGACUCC CCAAGGGGG UAAGAAACA  
6490 6500 6510 6520 6530 6540  
GCUUGCCUA UCGUUAACC UGACUUGGC GUCGCGUCU GCGAGAAAU GCGCCUUAU  
6550 6560 6570 6580 6590 6600  
GACAUACAC AAAAGCUUC UCAGGCGUA AUGGAGCUU CCAUGGCUU CCAUACUCC  
6610 6620 6630 6640 6650 6660  
CCGCGCAC CCGUGAGUA UCCUUGGAA GCAUGGGGG AAAAGAGGA CCGCAUGGU  
6670 6680 6690 6700 6710 6720  
UUUUGUAUG AUACCGAUG CUGGACUA ACCGUCACU AEAGAGACU CAGGACCGAG  
6730 6740 6750 6760 6770 6780  
GAGUUAUAU ACCAGGCUU CUCCUGGCC GAGGAGGCC GCAUGGCUU ACACUGGUC  
6790 6800 6810 6820 6830 6840  
ACUGAGAGC UUUACGAGG AGGGCCAGG UUCAACAGA AGGCUACAC CUGCGGUAC  
6850 6860 6870 6880 6890 6900  
AGACGUGGC GCGCAGCGG GCGGCUAAC ACUAGCAUG GUAAACCAU CACAUACUA

【図 2 F】

6910	6920	6930	6940	6950	6960
GUGAAAGCCC	UAGGGGCCUG	CAAGGCUGCG	GGGAUAGUUG	GGCCACAAU	GCUGGUAUCC
6970	6980	6990	7000	7010	7020
GGCGAUGACC	UAGUAGUCAU	CUCAGAAAGC	CAGGGGACUG	AGGAGGAOGA	GCGGAACCTG
7030	7040	7050	7060	7070	7080
AGAGCCUACA	CGGAGGCCAU	GACCAGGUAC	UCUGCCCCUC	CUGGUGAUCC	CCCCAGACCG
7090	7100	7110	7120	7130	7140
GAAUAUGACC	UGGAGCUAAU	AACAUCUUGU	UCCUCAAAUG	UGUCUGUGGC	GULGGGCCCC
7150	7160	7170	7180	7190	7200
GGGGGCGGCC	GCAGAUACUA	CCUGAGCCAGA	GACCCAAACCA	CUCCACUCCG	CCGGGCGUGCC
7210	7220	7230	7240	7250	7260
UGGGAAACAG	UUAGACACUC	CCCUAUCAAU	UCAUGGCUUG	GAAACAUCAU	CCAGUAUGCU
7270	7280	7290	7300	7310	7320
CCAACCAUAU	GGGUUCGCAU	GGUCCUAAUG	ACACACTUUCU	UCUCCAUUCU	CAUGGUCCAA
7330	7340	7350	7360	7370	7380
GACACCCUGG	ACCAGAACCU	CAACUUUGAG	AUGUAUGGAU	CAGUAUACUC	CGUGAAUCCU
7390	7400	7410	7420	7430	7440
UUGGACCUUC	CAGCCAUAUU	UGAGAGGUUA	CACGGGCUUG	ACGCCUUUUC	UAGGCACACA
7450	7460	7470	7480	7490	7500
UACUCUCACC	ACGAACUGAC	GCGGGUGGCU	UCAGCCCUCA	GAAAACUUGG	GGCGCCACCC
7510	7520	7530	7540	7550	7560
CUCAGGGUGU	GGAGAGUUG	GGCUCGCGCA	GUCAGGGCGU	CCCUCAUCUC	CCGUGGAGGG
7570	7580	7590	7600	7610	7620
AAAGCGGCCG	UUUGCGGCCG	AUAUCUCUUC	AAUUGGGCGG	UGAAGACCAA	GCUCAAAACUC
7630	7640	7650	7660	7670	7680
ACUCCAUUGC	CGGAGCGCGG	CCUACUGGAC	UUAUCCAGUU	GGUUCACCGU	CGGCGCGGGC
7690	7700	7710	7720	7730	7740
GGGGGCGACA	UUUUUCACAG	CGUGUCCGCG	GCCCGACCCC	GCUCAUUACU	CUUGGGCCUA
7750	7760	7770	7780	7790	7800
CUCCUACUUU	UGUAGGGGU	AGGCCUCUUC	CUACUCCCG	CUCGUAGAG	CGGCACACAC
7810	7820	7830	7840	7850	7860
UAGGUACACU	CCAUAGCUAA	CGGUUCCUUU	UUUUUUUUUU	UUUUUUUUUU	UUUUUUUUUU
7870	7880	7890	7900	7910	7920
UUUUUUUUUU	UUUUUUUUUU	UUUUUCCUUC	UUUCUUCCCU	UCUCAUCUUA	UUUCUACUUU
7930	7940	7950	7960	7970	7980
UUUCUUGGUG	GCUCCAUCUU	AGCCCUAGUC	ACGGCUAGCU	GUGAAAGGUC	CGUGAGCCCG
7990	8000	8010	8020	8030	8040
AUGACUGCAG	AGAGUGCCGU	AACUGGUCUC	UCUGCAGAUC	AUGU	



【図 3 A】

10 AACCGGCCU 20 AAUAGGGGG 30 ACAGUCCGCC 40 AUGAAUCACU 50 CCCCUGUGAG 60 GAACUACUGU  
70 CUUACGCGAG 80 AAAGCGUCUA 90 GCCAUGGGU 100 UAGUAGAGU 110 GUUGUACAGC 120 CUCCAGGCC  
130 CCCCCCGCG 140 GGAGAGCCAU 150 AGUGGUCGCG 160 GGAACCGGUG 170 AGUACACCGG 180 AAUUGCCGGG  
190 AAGACUCCGU 200 CCUUCUUGG 210 AUAAACCCAC 220 UCUAUGCCCG 230 GCCAUTUGGG 240 CGUGCCCCCG  
250 CAAGACUGCU 260 AGCCGAGUAG 270 CGUGGGGUG 280 CGAAGGGCCU 290 UGUGGUACCG 300 CCUGAUAGGG  
310 UGCUUGCGAG 320 UGCCCCGGGA 330 GGUCUCGUAG 340 AOCUGCACC 350 AUGAGCACAA 360 AUCCCAAACC  
370 UCAAGAAAA 380 ACCAAAAGAA 390 ACACUAAACG 400 UCGCCCAAG 410 AUUGAACAG 420 AUGGAUUGCA  
430 CGCAGGUUCU 440 CCGGCCCUU 450 GGGUGGAGAG 460 GCUAUUCCG 470 UAUSACUGG 480 CACAACAGAC  
490 AAUCGCGUGC 500 UCUGAUCCG 510 CCGUGUCCG 520 GCUGUCAGCG 530 CAGGGCGCC 540 CGGUUCUUU  
550 UGUCAGAGCC 560 GACCUGUCCG 570 GUGCCUGAA 580 UGAACUGCAG 590 GAAGAGGCG 600 CGCGCCUAC  
610 GUGGCGGCC 620 ACGACCGCG 630 UUCUUGGCG 640 AGCUGGCGC 650 GAUGUUGCA 660 CCGAAGCGG  
670 AAGGAGCGG 680 CCGCUAUGG 690 GCGAGUGCC 700 GGGGCGAGU 710 CUCCUGUACU 720 CUCAUCCUG  
730 UCCUGCGAG 740 AAAGUAUCCA 750 UCAUGGCGA 760 UGCAGUGCG 770 CGGCGCAUA 780 CGCUUGAUCC  
790 GGCUAUCCG 800 CCUUCGACC 810 ACCAAGCGA 820 ACAUCCGAC 830 GAGCGAGCAC 840 GUACUCGGAU  
850 GGAAGCGGU 860 CUUGUGAUC 870 AGGAUGAUC 880 GGAAGAAGAG 890 CAUCAGGGG 900 UCGCGCCAGC  
910 CGAAUUGUUC 920 GCGAGGCUA 930 AGGCGCGCU 940 GCGCGAGCG 950 GAGGAUCCG 960 UCGTACCCCA  
970 UGGCGAUGC 980 UGCUUGCGA 990 AUUUCAGGU 1000 GGAAAUUGG 1010 CGCUUUCUG 1020 GAUUCACGA  
1030 CUGUGCCCG 1040 CUGGGUGUG 1050 CGGACGCUA 1060 UCAGGACUA 1070 GCGUUGGUA 1080 CCGUGAUAU  
1090 UCCGAGAGG 1100 CUUGCGGCG 1110 AAUGGCGUA 1120 CCGCUUCCG 1130 GUGCUUACG 1140 GUAUCCCGC  
1150 UCCCGAUUG 1160 CAGCGCAUG 1170 CCUUCUAUG 1180 CCUUCUUGAC 1190 GAGUUCUUC 1200 GAGUUUAAC  
1210 CCUCUCCUC 1220 CCCCCCCC 1230 AACGUUACG 1240 GCGAAGCCG 1250 CUUGGAUAA 1260 GCGCGGUGG  
1270 CGUUGUCUA 1280 DAUGUUAUU 1290 UCCACCAUA 1300 UCGCGUCUU 1310 UGGCAAGUG 1320 AGGCGCCGA  
1330 AACUUGGCC 1340 UGUUUCUUG 1350 ACGAGCAUUC 1360 GAGGGGUCU 1370 UCCCCUCUC 1380 GCGAAGGGA

【図 3 B】

1390	1400	1410	1420	1430	1440
UGCAAGGUCU	GUUGAAUUC	GUAGAGGAG	CAGUUCUCU	GGAAGCUUCU	UGAAGACAAA
1450	1460	1470	1480	1490	1500
CACAGUCUGU	AGCGACCCUU	UGCAGGCAGC	GGAAACCCCC	ACCUAGGCGAC	AGGUGCCUCU
1510	1520	1530	1540	1550	1560
GCGGCACAAA	GCCACGUGUA	UAGAUACAC	CUGCAAGGC	GGCACAAACC	CAGUGCCACG
1570	1580	1590	1600	1610	1620
UUGUCAGUUG	GAUAGUGUUG	GAAAGAGUCA	AUUGGCUUCU	CUCAGCGUA	UUCACAAAGG
1630	1640	1650	1660	1670	1680
GGCUGAAGSA	UGCCACAGAG	GUACCCCAU	GUUUGGGAUC	UGAUCUGGG	CCUCGGUGCA
1690	1700	1710	1720	1730	1740
CAUGCUUAC	AUGGCUUAG	UCGAGGUUAA	AAAAAGUCU	AGGCCCCCG	AACCACGGGG
1750	1760	1770	1780	1790	1800
ACGUGGUUU	CCUUAGAAA	ACACGAUAAU	ACCAUGGCCC	CCAUCAACGC	UUACGCCACG
1810	1820	1830	1840	1850	1860
CAGACACAG	GUUCUUGGG	CUUUAUAGUG	GUGAGCAUGA	CGGGGCGUGA	CAAGACAGAA
1870	1880	1890	1900	1910	1920
CAGCCCGGG	AGGUCCAGU	CCUCCCACA	GUCACUCAGU	CCUCCUUGG	AACAUCCAUU
1930	1940	1950	1960	1970	1980
UCGGGGGUCU	UAUGGACUGU	UUACCAAGGA	GCUGGCAACA	AGACACUAGC	CGGCUCCGGG
1990	2000	2010	2020	2030	2040
GGCCCGGUA	CCGAGAUUA	CUAGAGGCGC	GAGGGGACU	UGGUCGGGUG	GCCCAGCCCU
2050	2060	2070	2080	2090	2100
CCUGGAGCA	AAUUCUUGA	GCCGUGUACG	UGUGGAGCGG	UCGACCUUGA	UUUGGUCAAG
2110	2120	2130	2140	2150	2160
CGGAAGGUCU	AUGUCAUCC	GGCUCGAGA	CGCGGGGACA	ACCGGGGAGC	GCUCUCUCC
2170	2180	2190	2200	2210	2220
CCGAGACCC	UUGGACCUU	GAAGGGGUC	UCGGGGGAGC	CUGUGCUUUG	CCUAGGGGC
2230	2240	2250	2260	2270	2280
CACGCUUGUG	GAUCCUCCG	GGCAGCUGUG	UGCUCUUGGG	GUGUGGCUAA	GUCCAUAGAU
2290	2300	2310	2320	2330	2340
UUCAUCCCG	UUGAGACGCU	CGACAUUGUC	ACGCGGUCUC	CCACCUUAG	UGACACAGC
2350	2360	2370	2380	2390	2400
ACACACACG	GUGUGCCCA	GACCUAUCAG	GUGGGGUACU	UGCAGCCCC	CACUGGCAU
2410	2420	2430	2440	2450	2460
GGAAAAAGCA	CCAGGUCCC	CGUCCGUAU	GCCGCCCAGG	GGUAUAAGU	GUUGGUGCUC
2470	2480	2490	2500	2510	2520
AAUUCUUGG	UGGCUGCCAC	CCUGGGAUUU	GGGCGUACU	UGUCCAGGC	ACAUGGCAUC
2530	2540	2550	2560	2570	2580
AAUCCCAACA	UUGGAGCUGG	AGUCAGAACU	GUGACGACCG	GGGAGCCCAU	UACAUAUCC
2590	2600	2610	2620	2630	2640
ACGUAUGUA	AAUCCUUGC	CGAUGGGGC	UGGCGAGGCG	GCGCCUUGA	CAUCAUCAUA
2650	2660	2670	2680	2690	2700
UGGGAUAAU	CCACCUUGU	GAUGCUACC	ACUAUUCUUG	GCAUCGGGAC	AGUCCUAGAC
2710	2720	2730	2740	2750	2760
CAAGCAGAGA	CAGCCGGGU	CAGGCUAACU	GUACUGGCCA	CGGCCACCC	CCCGGGGUGG

【図 3 C】

2770	2780	2790	2800	2810	2820
GUGACAAACC	CCCAUCCCAA	UAUAGAGGAG	GUAGCOCUOG	GACAGGAGGG	UGAGAUCCCC
2830	2840	2850	2860	2870	2880
UUCUAUUGGA	GGGCGUUUCC	CCUGUCUUAU	AUCAAGGGAG	GGAGGCACUU	GAUUUUUCUG
2890	2900	2910	2920	2930	2940
CACUCAAAGA	AAAAGUGUGA	CGAGCUUGCA	ACGGCCCUUC	GGGGCAUGGG	CUUGABCGCU
2950	2960	2970	2980	2990	3000
GUGGCAUADU	ACAGAGGGUU	GGACGUCUCC	AUAAUACCAA	CUCAAGGAGA	UGUGGUGGUC
3010	3020	3030	3040	3050	3060
GUUGCCACCG	AUGCCCUCAU	GACGGGGUAD	ACUGGAGACU	UUGACUCUGU	GAUCCGACUC
3070	3080	3090	3100	3110	3120
AACGUAGCGG	UACCCACGGC	CGUAGACUUC	AGCCUGGACC	CCACCUUCAC	UAUAAACCA
3130	3140	3150	3160	3170	3180
CAGACUGUCC	CGCAGACGGC	UGUCUCUGGU	AGUCAGCGCC	GAGGGCGCAC	GGGUAGAGGA
3190	3200	3210	3220	3230	3240
AGACUGGGCA	UUUAUAGUA	UGUUUCCACU	GGUGAGCGAG	CCUCAGGAU	GUUUGACAGU
3250	3260	3270	3280	3290	3300
GUAGUACUCU	GUGAGUGCUA	CGACCGAGGA	GCUGCUUGGU	AUGAGCUCUC	ACCAUGUGAG
3310	3320	3330	3340	3350	3360
ACGACCGUCA	GGCUCAGGGC	GUUUUUAAC	ACGCGCGGCU	UGCCUGUGUG	CCAGGAACAC
3370	3380	3390	3400	3410	3420
CUUGAGUUUU	GGGAGGAGGU	UUUACCGGCG	CUACACACAC	UAGAGCGUCA	UUUUUUUUCC
3430	3440	3450	3460	3470	3480
CAGACAAAGC	AGUGGGGGGA	AAAUUUUGCA	UACUUAUAG	CCUAUCAGGC	CACAGUGUGC
3490	3500	3510	3520	3530	3540
GCCAGGGGCA	AAGCGCCCCC	CCCGUCCUGG	GACGUCAUGU	GGAAGUGCUU	GACUCGACCC
3550	3560	3570	3580	3590	3600
AAGCCACGCG	UUGUGGGCCC	UACACCUUUC	CUUGACCGUU	UGGGCUCUGU	UACCAACGAG
3610	3620	3630	3640	3650	3660
GUACCCUUUA	CACACCCCGU	GACAAAATAC	AUCGCCACAU	GCAUGCAAGC	UGACCUUGAG
3670	3680	3690	3700	3710	3720
GUCAUGACCA	GCAUGUGGGU	CCUGGCUGGG	GGAGUCUUAU	CAGCGGUUGC	CGCGUAUUGC
3730	3740	3750	3760	3770	3780
UUAGCGACCG	GGGUGUUUCC	CAUCAUUGGC	CGUUUACACA	UCAAACGAGC	AGCUGUCUGU
3790	3800	3810	3820	3830	3840
GCUCGGGACA	AGGAGGUCCU	CUAUCAGGCU	UUUGAUGAGA	UGGAGGAUUG	UGCCUCCAGA
3850	3860	3870	3880	3890	3900
GCGGCUUCCU	UUGAAGAGGG	GCAGCGGAUA	GCCGAGAUUC	UGAAGUCCAA	GAUCCAGGGC
3910	3920	3930	3940	3950	3960
UUAAUUGCAG	AAGCCUCUAA	ACAGGCCCCG	GACAUACAA	CCGCUUGUCA	AGCUUUGUGG
3970	3980	3990	4000	4010	4020
CCCAAGAUUG	AGCAAUUCUG	GGCCAAACAU	AUGUGGAACU	UCAAUAGGGG	CAUUCAGUAC
4030	4040	4050	4060	4070	4080
CUUGCAGGAC	UGUCACACAU	GCCAGGGAAC	CCUGCUUGGG	CUUCCAUGAU	GGCAUUCAGC
4090	4100	4110	4120	4130	4140
GCCGCCCUCA	CCAGUCUGUU	GUCAACUAGC	ACCACCAUCC	UUUUUAAACU	UCUGGGGGGC

【図 3 D】

4150 4160 4170 4180 4190 4200  
GEGCUGGGU CCCAAAUUGC GGCACCCGCG GGGGCCACUG GCUUUGUUGU CAGUGGCCUG

4210 4220 4230 4240 4250 4260  
GUGGGAGCUG CUGUUGGCAG CAUAGGCCUG GGUAAAGUGC UGGUGGACAU CCUGGCCAGG

4270 4280 4290 4300 4310 4320  
UAUGGUGGG GCAUUGCGG GGCOCUGUC GCGUUUAGA UCAUGUCUGG CGAGAAAGCC

4330 4340 4350 4360 4370 4380  
UCCAUGGAGG AUGUCAUCA CUUGCUGCCU GGAUUCUGU CUCCAGGUGC UCUGUGGUG

4390 4400 4410 4420 4430 4440  
GGAGCAUCU GCGCGGCCAU UCUGCGCGC CAUGUGGAC CGGGGAAGG CGCGGUCCA

4450 4460 4470 4480 4490 4500  
UGGAUGACA GGCUUAUCC CUUCGCUCC AGAGGAAACC ACGUCGCCC UACUCACUAC

4510 4520 4530 4540 4550 4560  
GCGAGGAGU CGGAUGGUC GCAGGUGUC ACCCAAGUC UUGGCUUCU CACUAUAAU

4570 4580 4590 4600 4610 4620  
AGUCUACUA GGAGCUUCA CAACUGGUC ACUGAGGAU GCGCAUCC AUGGCGGCG

4630 4640 4650 4660 4670 4680  
UGUGGCUCC GCGAUGUG GGCUGGGUC UGUACCAUCC UAACAGAUU UAAGAACUGG

4690 4700 4710 4720 4730 4740  
CGAGCCUCCA AGCGGUCGC AAAGAUGCC GGCUCUCCU UUAUCUUGG CCAAAAGGG

4750 4760 4770 4780 4790 4800  
UACAAAGGCG UGUGGGCGG CACUGGCAC AUGACACAC GAUGGCGUG CGGGGCCAAC

4810 4820 4830 4840 4850 4860  
AUCUCUGCA ACGUCCCUU GGCUCUAG AGAUACAG GACCAARAC CUGCAUGAC

4870 4880 4890 4900 4910 4920  
ACUGGCGAG GGCUCUCC UAUCAUUGU UAUAAGAAG GCCAGGCUU GCGGAAACC

4930 4940 4950 4960 4970 4980  
GCGUUAACU UCAAGACGC CAUCUGGAG GUGGGGCGU CAGAGUAAG GGAAGUGAG

4990 5000 5010 5020 5030 5040  
CAGCAAGGCU CAUAUGCUA UAUAACAGG CUGACACUG ACAACUAAA AGUCCUUGC

5050 5060 5070 5080 5090 5100  
CAACUCCCU CUCCAGAGU UUUUCUUGG GUGGAGGAG UAUAUUUA UAGGUCCGC

5110 5120 5130 5140 5150 5160  
CCACACCAA AGCGUUUUU CCGGAGGAG GUCUGUUA GCGUUGGCU CAUUUCAUU

5170 5180 5190 5200 5210 5220  
GUCGUGGGU CUACGCUUC CUGGACCCU GAGCCGACA CUGAGGUAG GAUUCUAG

5230 5240 5250 5260 5270 5280  
CUAACAGAC CAUCCAUAU CAGCGGAG GCUCCAGGC GCGUUUAGC GCGGGGUA

5290 5300 5310 5320 5330 5340  
CCCCAUUGG AGGCAAGCUC CUACAGGAG CAGCUGUGG CGCAUCGCU GCGAGCACC

5350 5360 5370 5380 5390 5400  
UGCACACCC ACGUAAGAC CUAUGAUG GACAUGGUG AUGCCAAAU GUUCUUGGG

5410 5420 5430 5440 5450 5460  
GGCGGCUA UACGGAAGA GUCUGAGUC AAAGUGUUG UUCUGGACU CCUGACUA

5470 5480 5490 5500 5510 5520  
AUGACGAGG AAGAGGCGA CCUGAGCCU UCAGUACAU CGAGUAUAD GUCCCCAGG

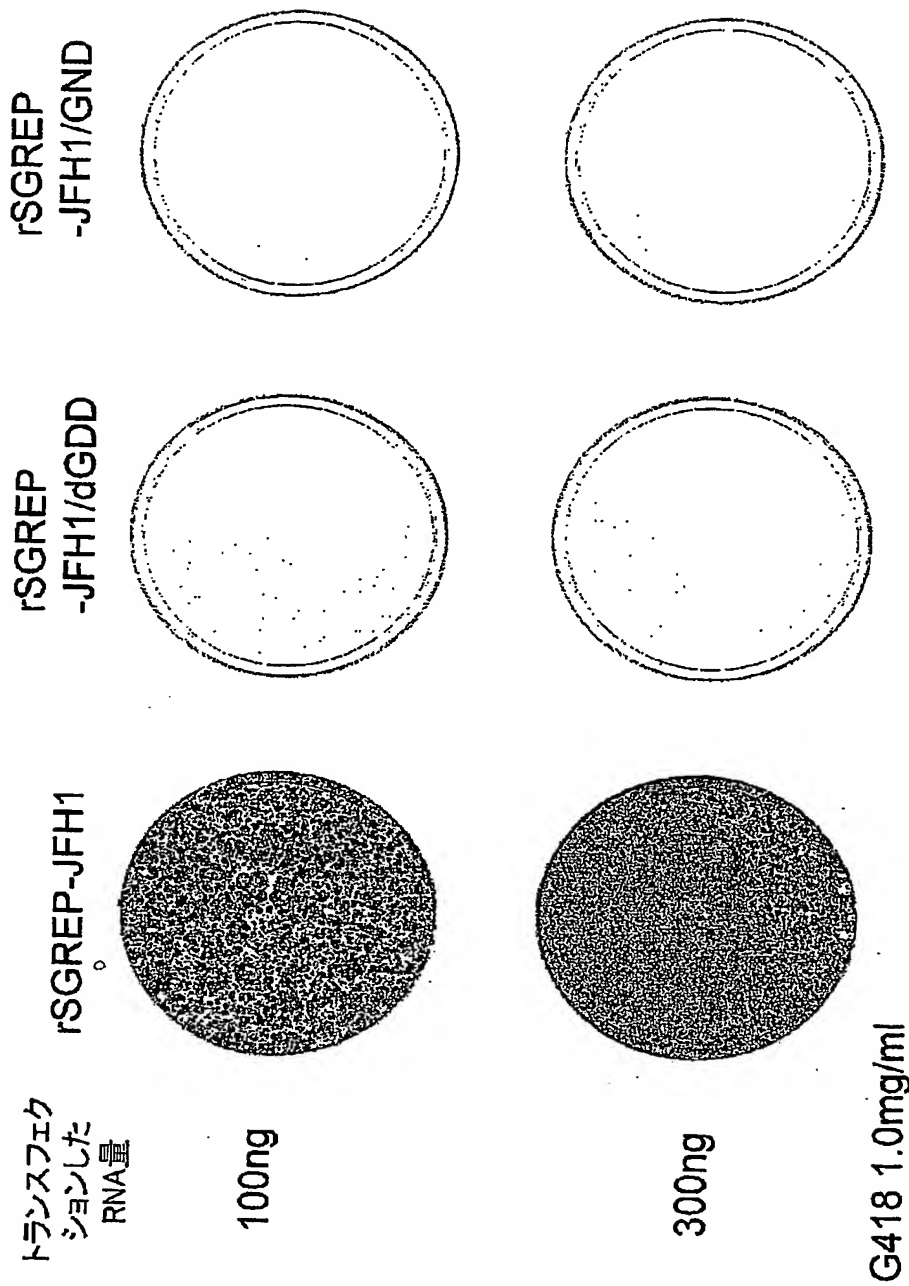
【図 3 E】

5530	5540	5550	5560	5570	5580
AAGAGGUGCC	CACCGGCTUU	ACCGGCTUAG	GCGCGGCCUG	AUUACAAACC	ACCGCUUSUG
5590	5600	5610	5620	5630	5640
GAUUGUGGA	AGAGGCCAGA	UUACCAACCA	CCACUGUUG	CGSGCGGGGC	UCGCCCCCCC
5650	5660	5670	5680	5690	5700
CCCAAAAGA	CCCGAGGCC	UCCGCCAAGG	AGACGCGGA	CAGUGGGUCC	GAGCGAGAGC
5710	5720	5730	5740	5750	5760
ACCAUAGGAG	AUGCCCUCCA	ACAGCUGGCC	AUCAAGUCCU	UUGGCCAGCC	CCGCCCAAGC
5770	5780	5790	5800	5810	5820
GGCGAUUCAG	GCCUUCAC	GGGGCGGAC	GCGCGCGACU	CGGGCGAUGG	GACACCCCUU
5830	5840	5850	5860	5870	5880
GACGAGUUGG	CUCUUUGGA	GACAGGUUCU	ACCUCCUCCA	UGCCCCCUU	CGAGGGGAG
5890	5900	5910	5920	5930	5940
CCUGGGGACC	CAGACCUUGA	GCCUGAGCAG	GUAGAGCUUC	AACCUCCUCC	CCAGGGGGG
5950	5960	5970	5980	5990	6000
GAGGCGGUC	CGGCUUGA	CUAGGGGUCC	UGGUCUACUU	CCUCCGAGGA	GGAUGACUCC
6010	6020	6030	6040	6050	6060
GUUGUGUGCU	GCUCCAUUUC	AUAUUCUUG	ACCGGGGCUU	UAUAUACUCC	UUGUGGCCCC
6070	6080	6090	6100	6110	6120
GAAGAGGABA	AGUGGCCAAU	UAAUCCUUG	AGCAACUCCG	UGUUGCGAUA	CCAUAAACAG
6130	6140	6150	6160	6170	6180
GUUACUGUA	CUACUCAA	GAGUGCCUCA	CUAAGGCUA	AAAAGGUAC	UUUUAUUGG
6190	6200	6210	6220	6230	6240
AUGCAAGGC	UOGAGCCUA	UUUAUGAUCA	GUUUAAAGG	ACAUCAGACU	AGGGGCCUCC
6250	6260	6270	6280	6290	6300
AAGGUCAGCG	CAGGCUCCU	CACCUAGAG	GAGGCGUGCC	AAUUGACCCC	ACCCACUCCU
6310	6320	6330	6340	6350	6360
GCAAGAUCCA	AGUAGGGUU	UGGGGCUAAG	GAGGUCGCA	GCUUGUCGG	GAGGGCCGUC
6370	6380	6390	6400	6410	6420
AAACACAUCA	AGUCGUGUG	GAAGGACUUC	UUGGAAGACU	CACAAACACC	AAUCCUACA
6430	6440	6450	6460	6470	6480
ACCAUCAGG	CCAAAAGGA	GGUGUUCUG	GUGGACCCC	CCAGGGGGG	UAAAAACCA
6490	6500	6510	6520	6530	6540
GCUOGCCUA	UGUUUACCC	UGAUCUGGC	GUAGGGUUCU	GCGAGAAGAU	GGCCUUUAU
6550	6560	6570	6580	6590	6600
GAUGUACAC	AAAAGCUCC	UCAGGCGGUG	AUGGGGCUU	CUUAGGGCU	CCAGUACUCC
6610	6620	6630	6640	6650	6660
CCCGCAGC	GGUGGGGUU	UCUCUUGAG	GCAUGGGGG	AAAAGAGGA	CCUUAUGGUU
6670	6680	6690	6700	6710	6720
UUUUCUAUG	AUAACCGAUG	CUUUGACUCA	ACCGUACUG	AGAGAGACAU	CAGGACUGAG
6730	6740	6750	6760	6770	6780
GAGUCCAUBU	ACCAGGCCUG	CUCCUUAACC	GAGGAGGCC	GAACUGCCAU	ACACUCGUG
6790	6800	6810	6820	6830	6840
ACUGAGAGAC	UCUAUGUGG	AGGGCCCAUG	UUCACAGCA	AGGGCCAGUC	CUGGGGUAC
6850	6860	6870	6880	6890	6900
AGUGUUGCC	GCGCCAGCG	GGUGCUUACC	ACUGUUGG	GGAACAACAU	CACAUGCUAU

【図 3 F】

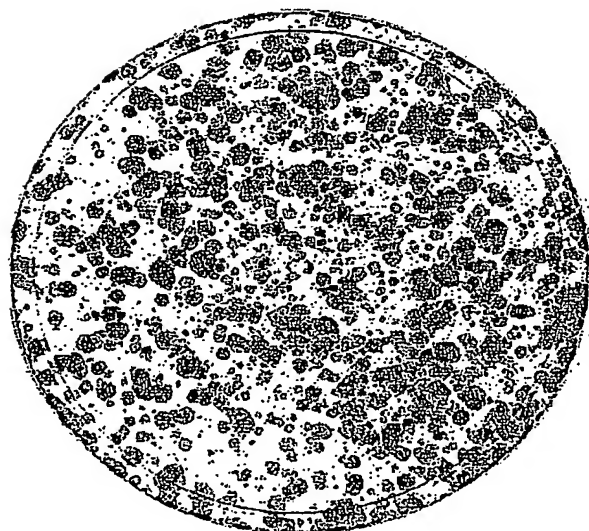
6910	6920	6930	6940	6950	6960
GUAAAAGCCC	UAGCGGCUUG	CAAGGCTUGG	GGGAUAAUUG	OGCCCACGAT	GCUGGUUAGC
6970	6980	6990	7000	7010	7020
GGGACGACU	UGGUUGUCAU	CUCAGBAAGC	CAGGGGACUG	AGGAGGACGA	GCGGAACCUG
7030	7040	7050	7060	7070	7080
AGAGCCUUCA	CGGAGGCUAU	GACCAGGUAU	UCUGCCCCUC	CUGGUGACCC	CCCCAGACCG
7090	7100	7110	7120	7130	7140
GAAUAUGACC	UGGAGCUAAU	AACAUCUUGU	UCCUCAAAAG	UGUCUGUGGC	ACUUGGCCCA
7150	7160	7170	7180	7190	7200
CAGGGCGGCC	GCAGAUACUA	CCUGACCAGA	GACCCACCCA	CUUCAAUUGC	CGGGGCUGCC
7210	7220	7230	7240	7250	7260
UGGGAACAG	UUAGACACUC	CCCUGUCAAU	UCAUGGCTUG	GAAACAUCAU	CCAGUACGCU
7270	7280	7290	7300	7310	7320
CCAACCAUUA	GGGUUCGCAU	GGUCCUGAUG	ACACACUUCU	UCUCCAUCUC	CAUGGCCCCAG
7330	7340	7350	7360	7370	7380
GACACCCUAG	ACCAGAACCU	UAACUUDGAA	AUGUACGGAU	CGGUGUACUC	CGUGAGUCCU
7390	7400	7410	7420	7430	7440
CUGGACCUC	CAGCCAUAAU	UGAAAGGUUA	CACGGGCUUG	ACGCCUUCUC	UCUGCACACA
7450	7460	7470	7480	7490	7500
UACACUCCCC	ACGAACUGAC	GCGGGUGGCU	UCAGCCCUCA	GAAACUUGG	GGCGCCACCC
7510	7520	7530	7540	7550	7560
CUCAGAGCGU	GGAGAGUUG	GGCGCGUGCA	GUUAGGGCGU	CCUCUACUC	CGUGGGGGG
7570	7580	7590	7600	7610	7620
AGGGGGGCGG	UUUGCGGUGG	GUACCUCUC	AACUGGGCGG	UGAGGACCAA	GCUCAAACUC
7630	7640	7650	7660	7670	7680
ACTCCUUGC	CGGAGGCACG	CCUCCUGGAU	UUGUCCAGUU	GGUUUACCGU	CGGCGCCGCG
7690	7700	7710	7720	7730	7740
GGGGCGACA	UUUADACAG	CGUGUCGCGU	GCCCGACCCC	GCCUAAUACU	CCUAGOCUA
7750	7760	7770	7780	7790	7800
CUCCUACUUU	CUGUAGGGGU	AGGCCUCUUC	CUACCCCCCG	CUCGAUAGAG	CGGCACACAU
7810	7820	7830	7840	7850	7860
UAGCUACACU	CCAUAGCUAA	CUGUCCUUUU	UUUUUUUUUU	UUUUUUUUUU	UUUUUUUUUU
7870	7880	7890	7900	7910	7920
UUUUUUUUUU	CUUUUUUUUU	UUUUUCCUC	UUUCUCCCU	UUCUACUUA	UUUACUUC
7930	7940	7950	7960	7970	7980
UUUCUUGGCG	GCUCCAUUU	AGCCCUAGUC	ACGGCUAGCU	GUGAAAGGUC	CGUGAGCCCG
7990	8000	8010	8020	8030	8040
AUGACUGCAG	AGAGUGCCGU	AACUGGUCUC	UCUGCAGAU	AUGU	

【図 4】

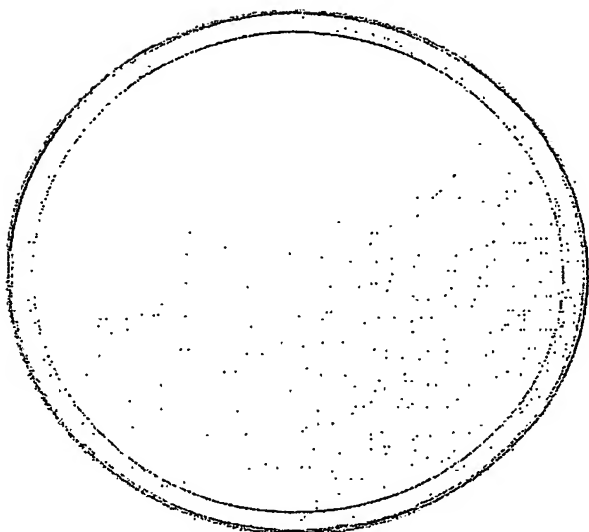


【図5】

rSGREP-JFH1



rSGREP-JCH1

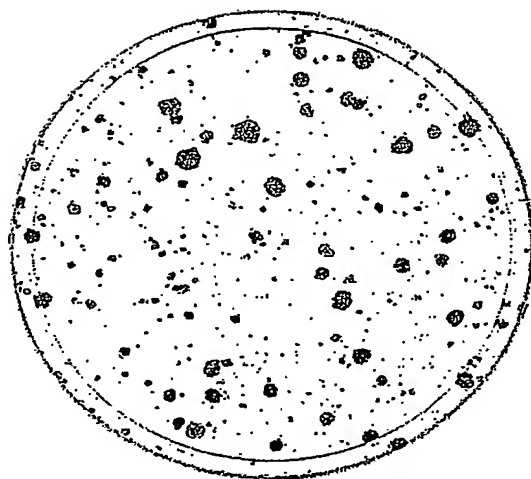


G418 0.5mg/ml  
トランスフェクションした  
RNA量 100ng

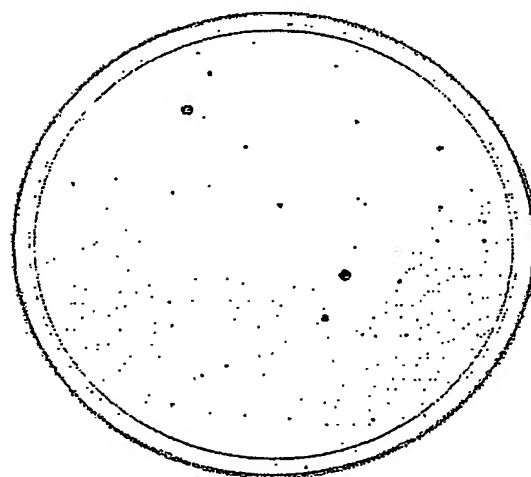


【図 6】

Mung Bean Nuclease  
処理あり



Mung Bean Nuclease  
処理なし

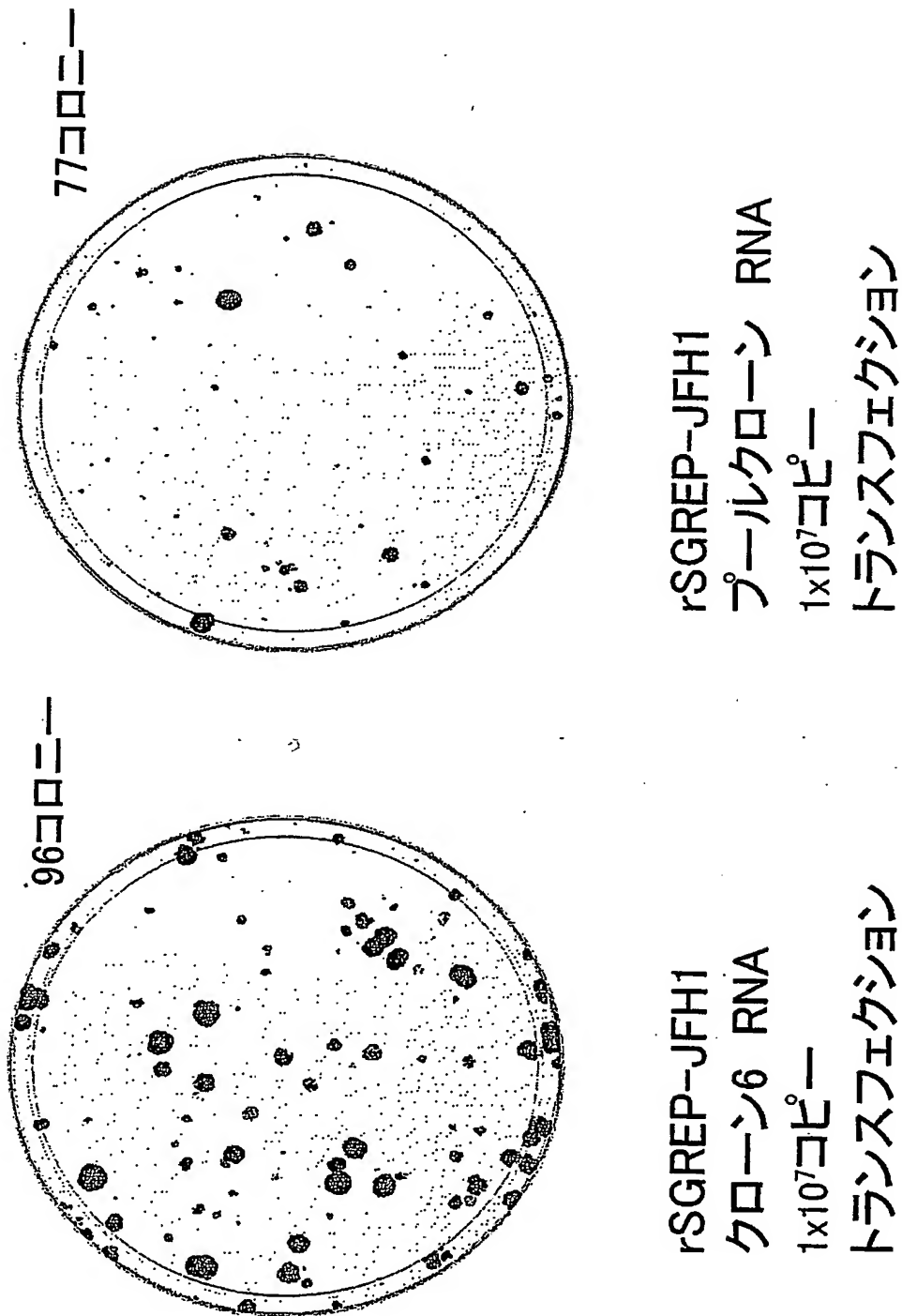


rSGREP-JFH1

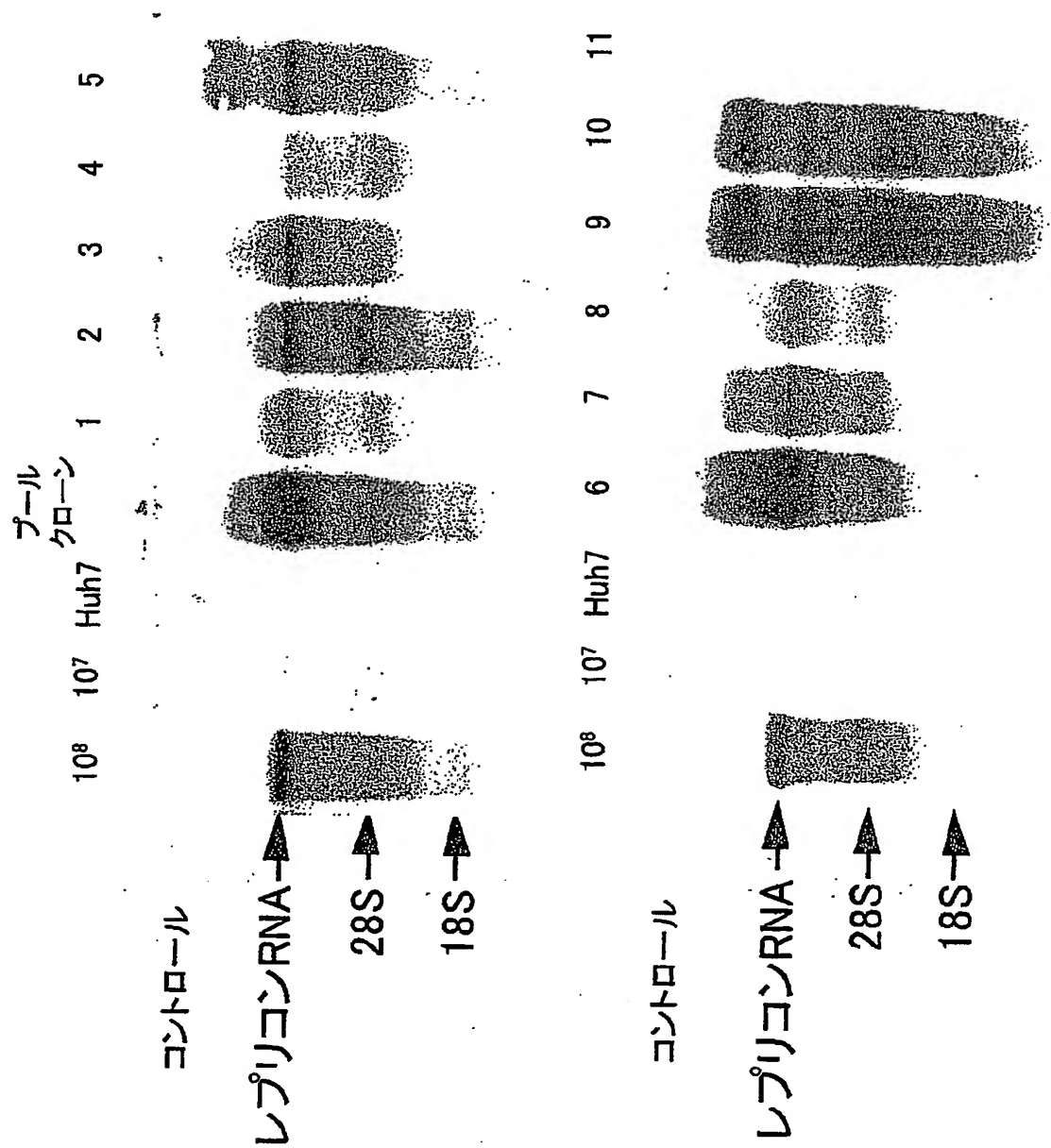
100ng

G418 1.0mg/ml

【図 7】



【図 8】

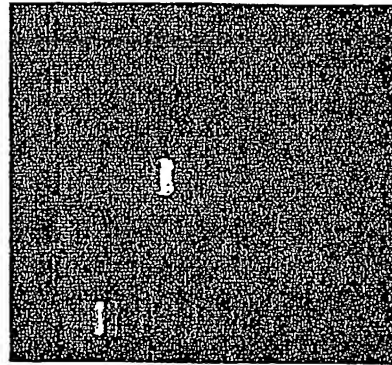


【図 9】

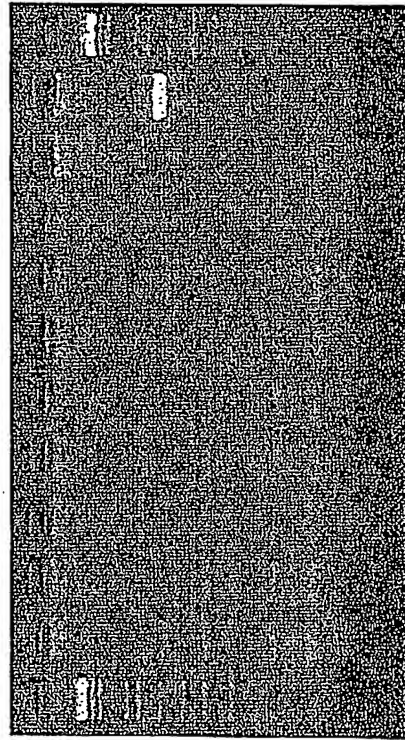
rSGREP-JCH1

rSGREP-JFH1

M 1 2 3 4 5 6



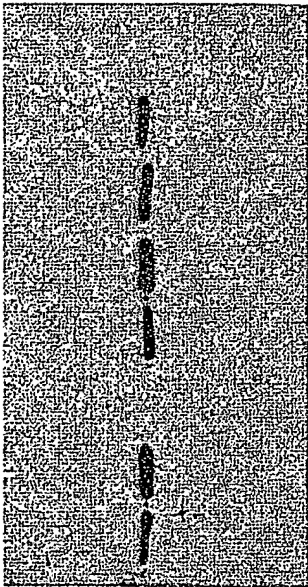
M 1 2 3 4 5 6 7 8 N P M



【図 10】

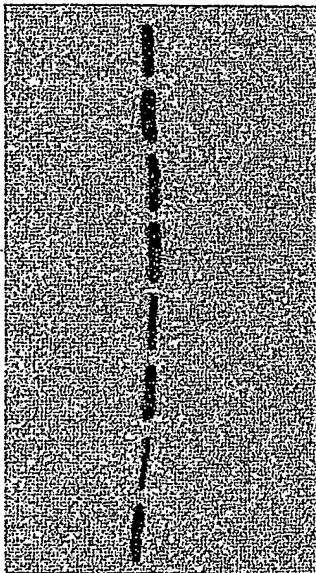
rSGREP-JCH1

1 2 3 4 5 6 P N

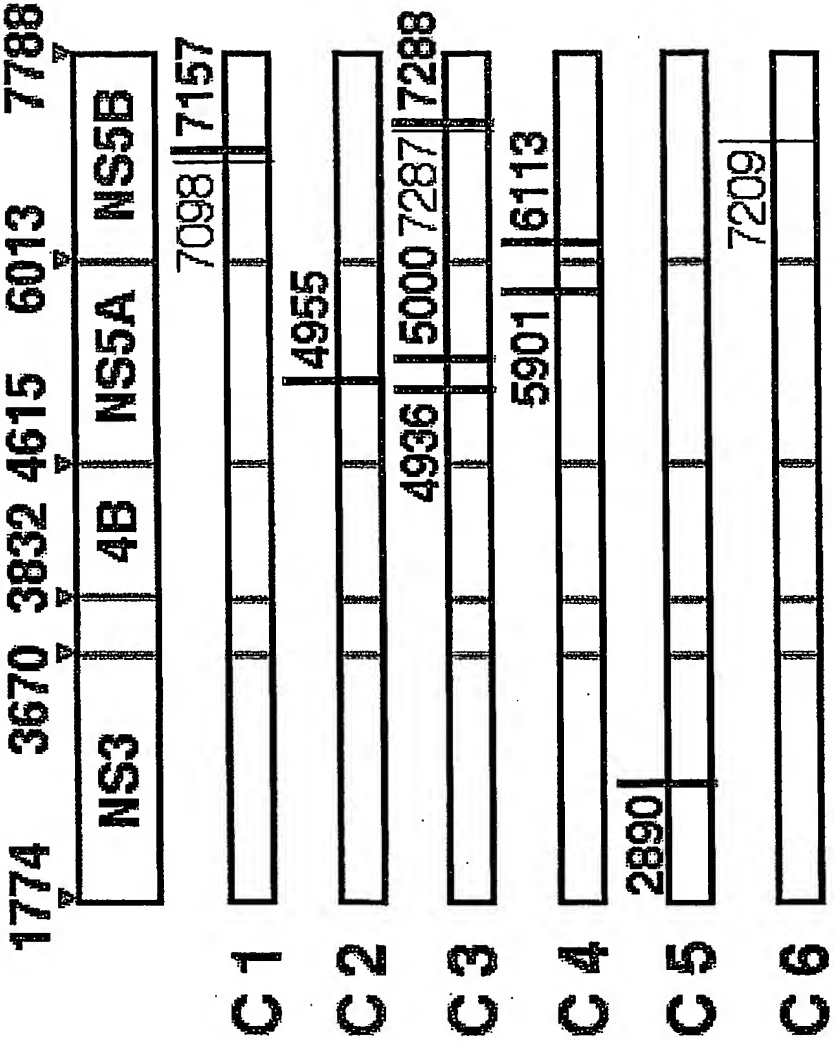


rSGREP-JFH1

1 2 3 4 5 6 7 8



【図 11】



【書類名】 要約書

【要約】

【課題】 遺伝子型1bとは異なる遺伝子型のHCV由来のレプリコンRNAを提供すること。

【解決手段】 遺伝子型2aのC型肝炎ウイルスのゲノムRNA上の、5'非翻訳領域と、NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列と、3'非翻訳領域とを少なくとも含む塩基配列からなる、レプリコンRNA。

【選択図】 なし

特願 2003-148242

出願人履歴情報

識別番号

[000003159]

1. 変更年月日

2002年10月25日

[変更理由]

住所変更

住 所

東京都中央区日本橋室町2丁目2番1号

氏 名

東レ株式会社



特願2003-148242

出願人履歷情報

識別番号

[591063394]

1. 変更年月日  
[変更理由]

1999年10月26日

名称変更

住所変更

住 所

東京都新宿区河田町10番10号

氏 名

財団法人 東京都医学研究機構

2. 変更年月日  
[変更理由]

2001年10月 9日

住所変更

住 所

東京都新宿区西新宿二丁目8番1号

氏 名

財団法人 東京都医学研究機構

特願 2003-148242

出 願 人 履 歴 情 報

識別番号

[503189262]

1. 変更年月日

2003年 5月26日

[変更理由]

新規登録

住 所

ドイツ連邦共和国 55099 マインツ

氏 名

ヨハネス グーテンベルク ユニベルスィテート マインツ